D

CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. III · NO. 3
1951

Published for

THE CANADIAN MATHEMATICAL CONGRESS

by the

University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, H. S. M. Coxeter, University of Toronto. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is \$6.00. This is reduced to \$3.00 for individual members of the following Societies:

Canadian Mathematical Congress
American Mathematical Society
Mathematical Association of America
London Mathematical Society
Société Mathématique de France

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of British Columbia
Ecole Polytechnique
Loyola College
McGill University
Université de Montréal
Royal Military College
Carleton College
Université Laval
University of Manitoba
McMaster University
Queen's University
University of Toronto

National Research Council of Canada
and the
American Mathematical Society
(Marion Reilly Fund)

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

ry,

in,
tz,

he
per
ble

ng

ent
is
he

the
al:

ON EUCLID'S ALGORITHM IN CYCLIC FIELDS

H. HEILBRONN

1. Introduction. In two papers I have proved that there are only a finite number of quadratic fields [6] and of cyclic cubic fields [7] in which Euclid's algorithm (E.A.) holds. Davenport has shown by a different method that there are only a finite number of quadratic fields [1, 2], of non-totally real cubic fields [3, 4] and of totally complex quartic fields in which E.A. holds.

The object of this paper is to extend these results to cyclic fields of higher degree. I shall prove

THEOREM 1. *For every $k \geqslant 4$ there are only a finite number of cyclic fields K of degree k whose discriminant Δ is the power of a prime, in which* E.A. *holds.*

The methods employed in this paper could actually furnish a proof of a theorem dealing with a more general type of cyclic field. But the classical theory of abelian fields allows us to name a large number of cyclic fields in which the class-number is greater than 1, and in which therefore E.A. cannot hold. Since these results are difficult to find in the existing literature, they will be quoted and proved in some detail in this paper.

To begin with we recall the two different definitions of the class-number of an algebraic field. H is the number of classes of ideals in an algebraic field if two ideals are considered equivalent provided their quotient is a principal ideal generated by a totally positive number; h is the number of classes of ideals in an algebraic field if two ideals are considered equivalent provided their quotient is any principal ideal. It is clear that $H = h$ for complex abelian fields.

We denote by $w(N)$ the number of distinct rational primes dividing a rational integer $N \neq 0$.

We call a cyclic field K a field of type T_1 if it is the composite field of cyclic fields K_j of degrees k_j and discriminants Δ_j where any two k_j are relatively prime, where any two Δ_j are relatively prime, and where $w(\Delta_j) = 1$.

We call a cyclic field K of degree k a field of type T_2 if it is the composition field of a field K_1 of type T_1 of odd degree, and of a cyclic field K_2 of discriminant Δ_2 of degree 2^l of the following type: $w(\Delta_2) \leqslant k$ and the discriminant of the unique subfield of K_2 of degree 2^{l-1} is a power of a prime, if $l > 1$. (For the purpose of this definition K_1 or K_2 may be the field of rational numbers.) We can now formulate

THEOREM 2. $(k, H) > 1$ *for a cyclic field K not of type T_1.*

THEOREM 3. $h > 1$ *for a cyclic field K not of type T_2.*

Received January 5, 1950.

All results of this paper and of my two previous papers can be summarized as

THEOREM 4. *For each* $k \geqslant 2$ E.A. *holds only in a finite number of cyclic fields* K *of degree* k *and discriminant* Δ, *if only fields of the following types are considered.*

(1) k *a prime.*

(2) $w(k) = 1$, k *odd.*

(3) $w(k) = 1$, K *complex.*

(4) $w(\Delta) = 1$.

(5) $w(\Delta) > w(k)$, k *odd.*

(6) $w(\Delta) > w(k)$, K *complex.*

(7) $w(\Delta) \geqslant k + w(k)$.

(8) $w(\Delta^*) > 1$ *for the discriminant* Δ^* *of every non-rational subfield* K^* *of* K.

(9) k *odd,* K *not of type* T_1.

(10) K *complex and not of type* T_1.

(11) K *not of type* T_2.

Finally I should like to mention two types of cyclic fields for which E.A. may possibly hold in an infinity of cases.

(a) The real quartic field

$$\rho(\sqrt{\tfrac{1}{2}(5 + 5^{\frac{1}{2}})p})$$

of discriminant $125p^2$, where $p \equiv 3 \pmod{20}$ is a prime.

(b) The complex sextic field

$$\rho((e^{2\pi i/9} + e^{-2\pi i/9}),\ (-p)^{\frac{1}{2}})$$

of discriminant -3^8p^3, where $p \equiv 3 \pmod 4$ is a prime.

We establish the following conventions. Small italics except e, i, and o denote positive rational integers, d, p and q denote positive rational primes. K, K', K_j etc. denote abelian fields of degrees k, k', k_j etc. and discriminants Δ, Δ', Δ_j etc.

Only absolutely abelian fields will be considered in this paper.

2. Dirichlet characters and Abelian fields. Two Dirichlet characters $\chi(n) \pmod m$ and $\chi'(n) \pmod{m'}$ are said to belong to the same train if and only if $\chi(n) = \chi'(n)$ for all n with $(n, mm') = 1$. Then each train contains exactly one primitive character $\chi_0(n) \pmod f$; f is called the conductor of the train, and also the conductor of all characters in the train. The product of two trains is defined in the obvious way, and it is clear that the trains form an infinite abelian group with respect to multiplication.

If $\chi(n)$ is a character mod m and if

$$m = m_1m_2,\ (m_1, m_2) = 1,$$

then $\chi(n)$ can be written in the form

$$\chi(n) = \chi_1(n)\chi_2(n)$$

where $\chi_1(n)$ and $\chi_2(n)$ are uniquely determined characters mod m_1 and m_2 respectively. In particular, if $\chi(n)$ is primitive, then $\chi_1(n)$ and $\chi_2(n)$ are primitive.

The principal results of class-field theory can easily be expressed in the following way.

Between all finite groups $\mathfrak{G}$ of trains and all abelian fields K there is a one-one relation [5, Theorem 1] which satisfies the following conditions:

I. The group $\mathfrak{G}$ is isomorphic to the Galois group of the field K. [5, Theorem 2.]

II. A field K' contains a field K if and only if the corresponding group $\mathfrak{G}'$ contains the corresponding group $\mathfrak{G}$. [5, Theorem 10.]

III. $|\Delta|$ equals the product of the conductors of the trains in $\mathfrak{G}$. [5, Theorem 16.]

IV.
$$\zeta_K(s) = \prod_{\chi} L(s, \chi),$$

where $\zeta_K(s)$ denotes the Dedekind ζ-function of K, and where χ runs through the primitive characters of the trains in $\mathfrak{G}$. [5, Theorem 14.]

V. If $\Delta = \pm p^l$, p becomes in K the kth power of a self-conjugate prime ideal of the first order.

VI. If $(\Delta, p) = 1$, p^l is the norm of an integral ideal in K if and only if $\chi(p^l) = 1$ for all primitive characters of the trains in $\mathfrak{G}$.

VII. If $(\Delta, n) = 1$, n is the norm of an integral ideal in K if and only if in the canonical representation

$$n = p_1^{l_1} \dots p_s^{l_s}$$

each factor $p_j^{l_j}$ is the norm of an integral ideal in K.

VIII. If n is the norm of an integral ideal in K, then $\chi(n) \geqslant 0$ for all characters of the trains in $\mathfrak{G}$.

IX. If K' is an abelian extension of K of relative discriminant 1, then the class-number H of K is divisible by k'/k. More precisely, the class-group of K contains a subgroup whose quotient group is isomorphic to the Galois group of K' over K. [5, Theorems 2 and 16.]

In addition we require two lemmas about discriminants.

LEMMA 1. *If the fields K_1 and K_2 have discriminants Δ_1 and Δ_2 and if $(\Delta_1, \Delta_2) = 1$, then the composition field $K_0 = K_1, K_2$ has discriminant*

$$\Delta_0 = \Delta_1^{k_2}\Delta_2^{k_1}$$

and degree $k_0 = k_1k_2$. [8, Theorem 88.]

Lemma 2. *If K' is an abelian extension field over K, then*

$$|\Delta'| = |\Delta|^{k'/k}$$

if and only if K' has relative discriminant 1 *over K.* [8, Theorem 39.]

3. Proof of Theorem 2. Let $\chi(n)$ be the primitive character in one of the trains which generate the group $\mathfrak{G}$ corresponding to K, so that k is the order of $\chi(n)$. Then we can write

$$\chi(n) = \chi_1(n) \dots \chi_s(n),$$

where $\chi_1(n), \dots, \chi_s(n)$ are primitive characters mod $p_1^{l_1}, \dots, p_s^{l_s}$ respectively, all the p_j being distinct. Let $k_1, \dots, k_s$ denote the order of $\chi_1(n), \dots, \chi_s(n)$ respectively; then the smallest common multiple

$$[k_1, \dots, k_s] = k.$$

Let P_j denote the product of the conductors of the characters

$$\chi_j(n), \chi_j^2(n), \dots, \chi_j^{k_j-1}(n) \qquad (1 \leqslant j \leqslant s).$$

Then the product of the conductors of the characters

$$\chi(n), \chi^2(n), \dots, \chi^k(n)$$

equals

$$P_1^{k/k_1} \dots P_s^{k/k_s} = |\Delta|$$

by III.

Let us now consider the group $\mathfrak{G}'$ of all trains generated by

$$\chi_1(n), \dots, \chi_s(n).$$

$\mathfrak{G}'$ contains the train $\chi(n)$, and the order k' of $\mathfrak{G}'$ equals

$$k' = k_1 \dots k_s.$$

The product of the conductors of all trains in $\mathfrak{G}'$ equals

$$(P_1^{k/k_s} \dots P_s^{k/k_s})^{k_1 \cdots k_s} = |\Delta^{k'/k}| = |\Delta'|$$

where Δ' is the discriminant of the field K' corresponding to $\mathfrak{G}'$.

It follows by I, II, and Lemma 2 that K' is an extension field of relative discriminant 1 over K. Hence by IX

$$H \equiv 0 \pmod{k'/k}.$$

Hence, if $k' > k$, then $(k, H) > 1$. If $k' = k$, then any two of the numbers $k_1, \dots, k_s$ are relatively prime, and the field K is of type T_1. This proves Theorem 2.

4. Proof of Theorem 3. We prove first:

Lemma 3. *If K is complex, then $H = h$. If K is real, then $h > 1$ unless*

the class group of K (in the narrow sense) is the direct product of not more than $k - 1$ abelian groups of order 2.

Proof. The first part of the lemma is trivial.

If K is real, then -1 is a non-totally positive unit in K. Therefore the group of all numbers in K, which are products of a unit in K, and of a totally positive number in K, is a subgroup of the group of all numbers ($\neq 0$) in K of index $\leqslant 2^{k-1}$. More precisely, the quotient group is the direct product of at most $k - 1$ groups of order 2. Since this quotient group is isomorphic to the quotient group of the two class groups in K, the lemma follows.

Assuming the notation used in the proof of Theorem 2, it suffices by virtue of IX and Lemma 3, to prove that, if the Galois group of K' over K is the direct product of at most $k - 1$ groups of order 2, then K is of type T_2.

Let K_0 be the field of largest odd degree k_0 which is contained in K, and let K_e be the field of largest degree $k_e = 2^l$ which is contained in K. Then K_0 and K_e are uniquely determined and we have

$$K = K_0, K_e, \quad k = k_0 k_e.$$

Let K_ϵ be the unique subfield of K_e of degree $k = \frac{1}{2}k_e$. Then we have to prove

(i) K_0 is of type T_1.

(ii) $w(\Delta_e) \leqslant k$.

(iii) $w(\Delta_\epsilon) = 1$ if $k_\epsilon > 1$.

We construct the extension field K'_0 over K_0 by the same process which gave us the extension field K' over K. If K_0 were not of type T_1, then $k'_0/k_0 > 1$ and odd. Since K'_0 is a subfield of K', we should have

$$(k'/k'_0)\,(k'_0/k_0) = (k'/k)\,(k/k_0),$$

which is a contradiction, because each factor on the right is a power of 2. This proves (i).

Next we construct the extension field K'_e by the same process. Again K'_e is a subfield of K' and we have

$$(k'/k'_e)\,(k'_e/k_e) = (k'/k)\,(k/k_e).$$

Here

$$k/k_e \equiv 1 \pmod 2, \; 2^{k-1} \equiv 0 \pmod{k'/k}.$$

If $w(\Delta_e) > k$, then

$$2^k \equiv 0 \pmod{k'_e/k_e}$$

which gives a contradiction. This proves (ii).

Finally if $k_e \geqslant 4$, $w(\Delta_e) \geqslant 2$, then the absolute Galois group of K'_e would have a subgroup of type (4, 4) by virtue of I. *A fortiori* the absolute Galois group of K' would have a subgroup of type (4, 4). Since the absolute Galois group of K is cyclic, the Galois group of K' over K would contain an element of order 4, which contradicts our hypothesis. This proves (iii).

5. Conventions and notations. We start by proving

LEMMA 4. *If K is cyclic, $w(\Delta) = 1$, $|\Delta| \geqslant k^{2(k-1)}$, then*

$$|\Delta| = d^{k-1}, d \equiv 1 \pmod{k}.$$

Proof. Let $\chi(n)$ be the primitive character in one of the trains which generate the group $\mathfrak{G}$ corresponding to K. By I and III $\chi(n)$ is a primitive character mod d^a (say) of order k. Hence

$$k|\varphi(d^a) = d^{a-1}(d-1),$$

which means that either $d|k$ or $k/d-1$. In the latter case, if $a > 1$, we should have a number n such that

$$\chi(n) \neq 1, n \equiv 1 (\mathrm{mod}\ d^{a-1}).$$

For this value of n

$$n^d \equiv 1 (\mathrm{mod}\ d^a),$$
$$1 = \chi(n^d) = \chi^d(n) = \chi(n) \neq 1,$$

which is a contradiction. Hence $\chi(n)$ is a character mod d; $\chi^j(n)$ is *a fortiori* a character mod d for $1 < j < k$, and it follows from III that

$$|\Delta| = d^{k-1}.$$

If d/k, we proceed as follows. We assume that

$$k = d^b m, \quad (m, d) = 1.$$

Then, for $d > 2$, the group of all characters mod d^a is cyclic of order $\varphi(d^a)$. Hence the number of characters mod d^a of order k equals $\varphi(k)$ if $k/\varphi(d^a)$ and 0 otherwise. Hence there exists a primitive character mod d^a of order k if and only if

$$\varphi(d^a) \equiv 0 \pmod{k}, \quad \varphi(d^{a-1}) \not\equiv 0 \pmod{k}.$$

This implies

$$m|(d-1), \quad a = b + 1,$$
$$d^a \leqslant dk \leqslant k^2, |\Delta| \leqslant k^{2(k-1)}.$$

If $d = 2$, $d|k$, the argument is similar. We may assume at once that $a > 3$. Then the group of characters mod d^a is abelian of type $(2, 2^{a-2})$, and the number of characters of order $k = 2^b$ equals 3 if $b = 1$, 2^b if $2 \leqslant b \leqslant a-2$, 0 if $b > a-2$.

Hence there exists a primitive character mod 2^a of order k if and only if

$$b = a - 2.$$

This implies

$$d^a = d^2 k \leqslant k^3, \quad |\Delta| \leqslant k^{3(k-1)}.$$

For the rest of the paper excluding the last paragraph we assume $k \geqslant 4$, K cyclic; hence by virtue of Lemma 4

$$d \equiv 1 \pmod{k}, \; |\Delta| = d^{k-1}.$$

$\chi(n)$ is again the primitive character mod d of order k in a train which generates the group $\mathfrak{G}$ corresponding to K. $\chi(n)$ will now be fixed.

Let A_j denote the class of integers n for which

$$\chi(n) = e^{2\pi i j/k} \qquad (0 \leqslant j \leqslant k-1).$$

Let B denote the subclass of integers b in A_0 for which

$$b = b_1 b_2, \; (b_1, b_2) = 1$$

implies b_1 in A_0. Let C denote the sub-class of integers c in A_0 which can be decomposed in the form

$$c = c_1 c_2, \;\; (c_1, c_2) = 1, \;\; \chi(c_1) \neq 1.$$

Clearly every number in A_0 lies either in B or in C. It follows from VI and VII that a number n is norm of an integral ideal in K prime to d if and only if n lies in B.

Also $q_1 < q_2$ are the two smallest primes not in A_0 which do not equal d; and r is the smallest number in C which is prime to q_1 and which satisfies

(1) $$r \equiv -d \pmod{4} \qquad \text{if } q_1 = 2,$$

(2) $$r \equiv -d^2 - 1 \pmod{9} \quad \text{if } q_1 = 3.$$

For $q_1 \geqslant 5$ no additional condition is imposed upon r.

Let ϵ be a positive number which will be fixed later; it may be arbitrarily small. The constants involved in the symbols O and o will depend on k only. Unless the contrary is stated the symbol o will refer to the limit as $d \to \infty$. We put

$$x = [d^{\frac{1}{2}+\epsilon}], \;\; y = [d^{\frac{1}{4}+\epsilon}].$$

6. Further lemmas.

LEMMA 5. $\sum_{p \leqslant z} p^{-1} = \log \log z + \gamma + o(1)$ *as* $z \to \infty$, *where* γ *is an absolute constant.* [9, Theorem 7].

LEMMA 6. *For each non-principal character* $\chi(n)$ mod m

$$\sum_{n=1}^{z} \chi(n) = O\,(m^{\frac{1}{2}} \log m) \qquad [10].$$

LEMMA 7. $q_2 \leqslant y$ *if d is sufficiently large.*

Proof. We assume $q_2 > y$. Then all primes $\leqslant y$, with the possible exception of q_1, belong to A_0. Hence, if

$$n \leqslant x, \; (n, q_1) = 1, \; p | n, \; y < p,$$

then $\chi(n) = \chi(p)$ unless n is divisible by the product pp' of two primes in the interval $y < p \leqslant x$, $y < p' \leqslant x$.

Therefore we have for $1 \leqslant j \leqslant k-1$

$$\sum_{\substack{n=1 \\ (n,\, q_1)=1}}^{x} \chi^j(n) = \sum_{\substack{n=1 \\ (n,\, q_1)=1}}^{x} 1 + \sum_{\substack{n=1 \\ (n,\, q_1)=1}}^{x} (\chi^j(n) - 1)$$

$$= (1 - q_1^{-1})x + O(1) + \sum_{\substack{y < p \leqslant x \\ p \neq q_1}} (\chi^j(p) - 1) \sum_{\substack{m \leqslant x/p \\ (m,\, q_1)=1}} 1 + \sum_{p > y} \sum_{\substack{p' > y \\ pp' \leqslant x}} O(x/pp')$$

$$= (1 - q_1^{-1})x + O(1) + \sum_{y < p \leqslant x} \{(\chi^j(p) - 1)\,(1 - q_1^{-1})xp^{-1} + O(1)\}$$

$$+ O(xy^{-1}) + O(x(\sum_{y < p < d^{\frac{1}{4}+2\epsilon}} p^{-1})^2)$$

$$= (1 - q_1^{-1})x\{1 + \sum_{y < p \leqslant x} (\chi^j(p) - 1)p^{-1}\} + O(\pi(x))$$

$$+ O\left\{x\left(\log \frac{\frac{1}{4} + 2\epsilon}{\frac{1}{4} - 2\epsilon} + o(1)\right)^2\right\} \qquad \text{(Lemma 5)}$$

$$= (1 - q_1^{-1})x\{1 + \sum_{y < p \leqslant x} (\chi^j(p) - 1)p^{-1}\} + O(\epsilon^2 x) + o(x).$$

Applying Lemma 6, this gives, after division by $(1 - q_1^{-1})x$,

$$0 = O(\epsilon^2) + o(1) + 1 + \sum_{y < p \leqslant x} (\chi^j(p) - 1)p^{-1}.$$

Summing this over $j = 1, \ldots, k-1$ we obtain

$$0 \geqslant O(\epsilon^2) + o(1) + k - 1 - k \sum_{y < p \leqslant x} p^{-1}.$$

Hence

$$\sum_{y < p \leqslant x} p^{-1} \geqslant 1 - k^{-1} + O(\epsilon^2) + o(1).$$

But by Lemma 6

$$\sum_{y < p \leqslant x} p^{-1} = \log\log x - \log\log y + o(1)$$

$$= \log \frac{\frac{1}{2} + \epsilon}{\frac{1}{4} - \epsilon} + O(1) = \log 2 + O(\epsilon) + o(1).$$

Hence

$$\log 2 \geqslant 1 - k^{-1} + O(\epsilon),$$

which is not true if ϵ is sufficiently small. This proves the lemma.

From now on ϵ is fixed as a function of k.

LEMMA 8. *$q_1 r < d^{1-\epsilon}$, if d is sufficiently large.*

Proof. We assume that d is so large that Lemma 7 applies. If q_2 lies in A_j $(1 \leqslant j \leqslant k-1)$, we choose for u the smallest number in A_{k-j} which satisfies

$$\begin{aligned} &(u, q_1q_2) = 1, \\ &uq_2 \equiv -d \quad (\text{mod } 4) \text{ if } q_1 = 2, \\ &uq_2 \equiv -d^2-1 \quad (\text{mod } 9) \text{ if } q_1 = 3. \end{aligned}$$

If d is sufficiently large, it is easily deduced from Lemma 6 that

$$u < x.$$

(The detailed argument is explicitly developed in [7].) Since uq_2 lies in C it follows from the definition of r that

$$r \leqslant uq_2 < xq_2,$$

and by Lemma 7 that

$$q_1r < q_1(xq_2) < xq_2^2 \leqslant xy^2 \leqslant d^{1-\epsilon}.$$

LEMMA 9. *If* $q_1 \geqslant 5$, $s < q_1$, *we can find a prime* p_0 *such that*

$$(p_0, s) = 1, \; p_0 < q_1, \; p_0 \leqslant \log d$$

provided d *is sufficiently large* [7, Lemma 4].

LEMMA 10. *For sufficiently large* d *we can write*

$$d = sr + tq_1,$$

where s *in* B, $(t, q_1) = 1$.

Proof. We distinguish three cases.

First case. $q_1 = 2$. We have

$$d = r + 2t,$$

and it follows from (1) that t is odd.

Second case. $q_1 = 3$. Then we have with $s = 1$ or $s = 2$

$$d = sr + 3t.$$

Clearly s lies in B, since $q_1 = 3$ is the smallest positive integer not in A_0. If t were divisible by 3, we should have by (2)

$$\begin{aligned} sr \equiv -s(d^2+1) &\equiv d && (\text{mod } 9), \\ (\pm 2s - 1)d &\equiv s(d \pm 1)^2 && (\text{mod } 9), \\ (-4s^2+1)d^2 &\equiv s^2(d^2-1)^2 && (\text{mod } 9), \\ -4s^2+1 &\equiv 0 && (\text{mod } 9), \end{aligned}$$

which is not true for $s = 1$ or $s = 2$.

Third case. $q_1 \geqslant 5$. Again, by Lemma 8, we can find s and t such that

$$d = sr + tq_1, \quad s < q_1.$$

Clearly, s lies in B, as it is not divisible by a prime $\geqslant q_1$.

But q_1 may possibly divide t. If $q_1|t$, we use the prime p_0 of Lemma 9 and denote by n the smallest positive solution of the congruence

$$s + nq_1 \equiv 0 \pmod{p_0}.$$

Then

$$s + nq_1 < q_1 + (p_0 - 1)q_1 = p_0 q_1. \tag{3}$$

We consider the representation

$$d = (s + nq_1)r + (t - nr)q_1.$$

Since $n < q_1$, $t - nr$ is prime to q_1. Since by (3), Lemma 9 and Lemma 8, for sufficiently large d

$$(s + nq_1)r < p_0 q_1 r \leqslant (\log d)d^{1-\epsilon} < d,$$

it follows that

$$t - nr > 0.$$

Finally it follows from (3) and Lemma 9 that no prime $\geqslant q_1$ divides $s + nq_1$. Hence $s + nq_1$ lies in B, and our lemma is proved in all cases.

LEMMA 11. *If d is sufficiently large,*

$$d = c + g,$$

where c lies in C, and g does not lie in B.

Proof. We assume that d is so large that Lemma 10 applies, and put

$$c = sr, \quad g = tq_1.$$

Clearly g does not lie in B, since

$$g = tq_1, \ (t, q_1) = 1, \qquad q_1 \text{ not in } A_0.$$

Since r lies in C, we have a decomposition

$$r = r_1 r_2, \ (r_1, r_2) = 1,$$

where r_1 does not lie in A_0. It follows from the fundamental theorem of arithmetic that we have a decomposition of s such that

$$s = s_1 s_2, \ (s_1, s_2) = 1, \ (r_1, s_2) = (r_2, s_1) = 1.$$

Since s lies in B, s_1 lies in A_0. We have a decomposition

$$c = sr = (s_1 r_1)(s_2 r_2), \ (s_1 r_1, s_2 r_2) = 1,$$

where $s_1 r_1$ does not lie in A_0. Hence c, lying in A_0, lies in C.

7. Proof of Theorem 1. We assume that E.A. holds in K. Then, by condition V, there exists in K a self-conjugate principal prime ideal (δ) of norm d.

We assume that d is so large that Lemma 11 applies. Since c lies in A_0, the congruence

$$n^k \equiv c \pmod{d}$$

has a solution. Since E.A. holds in K, we can find an integer γ in K such that

$$n \equiv \gamma \pmod{\delta}, \qquad |N(\gamma)| < |N(\delta)| = d.$$

Since (δ) is self-conjugate, the congruence

$$n \equiv \gamma' \pmod{\delta}$$

holds for each conjugate γ' of γ. Multiplying these k congruences we obtain

$$c \equiv n^k \equiv N(\gamma) \pmod{\delta},$$
$$c \equiv N(\gamma) \pmod{d}.$$

Hence

$$\text{either } N(\gamma) = c \text{ or } N(\gamma) = c - d = -g.$$

This means that the norm of the ideal (γ) equals c or g, which is impossible by Lemma 11 and condition VII.

8. Proof of Theorem 4. We take each individual assertion in Theorem 4, starting from the end.

(11) follows from Theorem 3.

(10) follows from Theorem 2, since $H = h$ if K is complex.

(9) follows from Theorem 3, since for odd k a field of type T_2 is a field of type T_1.

(8) If k is divisible by an odd prime, K_0 is not of type T_1, and therefore K is not of type T_2. If $k = 2^l$, $l \geqslant 2$, the field K_e has discriminant Δ_e with $w(\Delta_e) > 1$, hence K is not of type T_2. If $k = 2$, the result follows from my first paper [6].

(7) If K were of Type T_2, then

$$w(\Delta_0) \leqslant w(k_0)$$

and

$$w(\Delta_e) \leqslant k \text{ for even } k.$$

Hence

$$w(\Delta) \leqslant w(\Delta_0) + w(\Delta_e) \leqslant \begin{cases} w(k_0) < k + w(k) & \text{for odd } k. \\ w(k_0) + k = k - 1 + w(k) & \text{for even } k. \end{cases}$$

(6) K is not of type T_1, and $h = H > 1$.

(5) Since for odd k a field of type T_2 is a field of Type T_1, K is not of type T_2.

(4) follows from Theorem 1 for $k \geqslant 4$, and from my older results if $k = 2$ or $k = 3$.

(3) follows from (4) and (6).

(2) follows from (4) and (5).

(1) follows from (2) if k is odd, and from my older results if $k = 2$.

References

[1] H. Davenport, *Indefinite binary quadratic forms, and Euclid's Algorithm in real quadratic fields*, Proc. London Math. Soc., in course of publication.

[2] ——— *Indefinite binary quadratic forms*, Quart. J. Math., Oxford Ser. (2), vol. 1 (1950), 54-62.

[3] ——— *Euclid's Algorithm in cubic fields of negative discriminant*, Acta Math., vol. 84 (1950), 159-179.

[4] ——— *Euclid's Algorithm in certain quartic fields*, Trans. Amer. Math. Soc., vol. 68 (1950), 508-532.

[5] H. Hasse, *Bericht über neuere Untersuchungen aus der Theorie der algebraischen Zahlkörper*, Jber. Deutsch. Math. Verein., vol. 35 (1926), 1-55.

[6] H. Heilbronn, *On Euclid's Algorithm in real quadratic fields*, Proc. Cambridge Phil. Soc., vol. 34 (1938), 521-526.

[7] ——— *On Euclid's Algorithm in self-conjugate cubic fields*, Proc. Cambridge Phil. Soc., vol. 46 (1950), 377-382.

[8] D. Hilbert, *Bericht über die Theorie der algebraischen Zahlkörper*, Jber. Deutsch. Math. Verein., vol. 4 (1897), 175-546.

[9] A. E. Ingham, *The distribution of prime numbers* (Cambridge, 1932).

[10] G. Pólya, *Über die Verteilung der quadratischen Reste und Nichtreste*, Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl. 1918, 21-29.

The Royal Fort, Bristol 8

TWO VERTEX-REGULAR POLYHEDRA

HUGH APSIMON

1. Introduction. The definition of a regular polyhedron may be enunciated as follows:

(α) A polyhedron is said to be regular if its faces are equal regular polygons, and its vertex figures are equal regular polygons.

In a recent note[1] I gave three examples of uniform non-regular polyhedra, which I called *facially-regular*, using the definition:

(β) A polyhedron is said to be facially-regular if it is uniform and all its faces are equal.

It is evident that this definition may be replaced by one similar to that for regular polyhedra given above, such as[2]:

(γ) A polyhedron is said to be facially-regular if its faces are equal regular polygons, and its vertex figures are equal polygons.

A comparison of the definitions (α) and (γ) suggests that it would be of interest to investigate polyhedra which differ only from regular polyhedra in that the condition for regularity on the faces is dropped in (α), instead of the condition for regularity on the vertex figures (which gives (γ)). I call such polyhedra *vertex-regular*, using the definition:

(δ) A polyhedron is said to be vertex-regular if its faces are equal polygons and its vertex figures are equal regular polygons;

and in this note I describe two such polyhedra.

It will be noticed that if we define the regularity of a polyhedron in the more normal way:

(ϵ) A polyhedron is said to be regular if it possesses two particular symmetries; one which cyclically permutes the vertices of any face c, and one which cyclically permutes the faces that meet at a vertex C, C being a vertex of c;

then facially-regular polyhedra possess the first of these symmetries, and vertex-regular polyhedra possess the second.

For convenience I use the notation a^b to denote a vertex-regular polyhedron whose faces are a-gons, b of which meet at each vertex; but the symbol does not necessarily define the polyhedron uniquely.

The polyhedra which I describe are in this notation 12^3 and 9^3.

2. The polyhedron 12^3. From the ordinary space filling $[4^4]$, remove cubes

Received March 20, 1950.

[1] Can. J. Math., vol. 2 (1950), 326.

[2] For the sake of simplicity in this and the next definition I consider a polyhedron to be such that every face is accessible to any other face by paths crossing from one face to another by the edge common to both.

in such a manner that, for those remaining, each vertex of each member is a vertex of just one other member, and no two members have an edge in common.

Consider each cube as composed of 27 congruent component cubes. Squash the pile so that each cube retains its size and orientation, but corner components which originally touched come into coincidence (Fig. 1(a)).

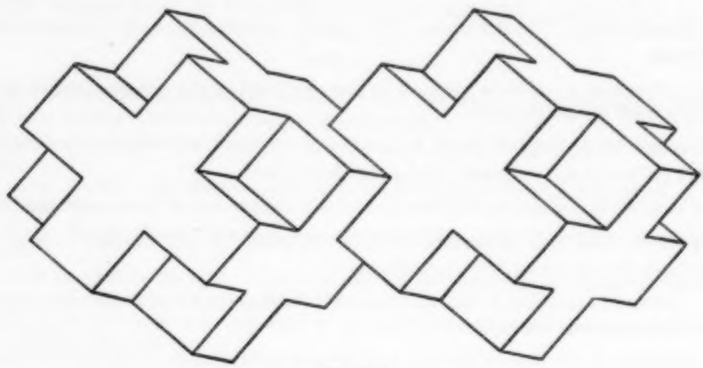

FIG. 1(a)

The surface of each cube not enclosed in the surface of any other cube now consists of six crosses (Fig. 1(b)), and the arrangement of all these crosses is such that each edge of each cross is in common with just one other cross. Further, three such crosses meet at each vertex, and their planes are mutually perpendicular, so that the vertex figure is an equilateral triangle. Hence the

FIG. 1(b)

polyhedron formed by the aggregate of the crosses has for its faces eqqal polygons and for its vertex figures equilateral triangles: it is vertex-regular.

The vertices of this polyhedron can best be described by referring it to rectangular Cartesian coordinates. Classify the integers by their residues (mod 8), and label P those points having odd coordinates one of which is congruent to 1 or 7, and another of which is congruent to 3 or 5 (the third being congruent to any one of 1, 3, 5, 7). Then the points P are the vertices of a 12^3 having edge 2, vertex figures equilateral triangles, and faces equal to the "Greek cross" whose vertices are in order:

$$\begin{array}{llllll} (3, 1) & (1, 1) & (1, 3) & (-1, 3) & (-1, 1) & (-3, 1) \\ (-3, -1) & (-1, -1) & (-1, -3) & (1, -3) & (1, -1) & (3, -1). \end{array}$$

The coordinates of the centres of the faces of this 12^3 lie at the points having coordinates congruent (mod 8) to one of the sets 4, 4, 1; 4, 4, 7; 0, 0, 3; 0, 0, 5; in some order. As I showed in my previous note, these points are the vertices of a facially-regular polyhedron which I described as 3^{12}, so that the polyhedra 3^{12} and 12^3 are reciprocal in the sense that the vertices of a 3^{12} lie at the centres of the faces of a 12^3 (but not *vice versa*).

3. The polyhedron 9^3. For clarity I first give a short description of the facially-regular polyhedron 3^9. Consider a set of octahedra each coloured α on two opposite faces and β on the remainder, and a set of tetrahedra coloured γ on all faces. Join these polyhedra by their faces according to the rule $\alpha \leftrightarrow \gamma$; then the β faces are the faces of a 3^9. It will be seen that one could alternatively consider the 3^9, viewed as a solid body, as composed of suitably selected tetrahedra and octahedra from the space filling of octahedra and tetrahedra[3] $[3^4]$; and that the faces of 3^9 lie in planes parallel to any of its component tetrahedra according to the system shown in Fig. 2(a).

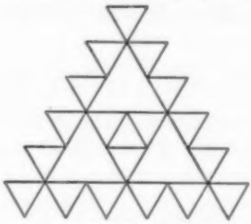

FIG. 2(a)

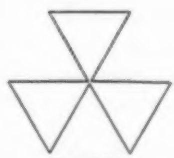

FIG. 2(b)

The faces of a 3^9 occur in sets of three. Regard each set of three as constituting an irregular enneagon, having as vertices the vertices of a regular hexagon and its centre (which is counted three times). Such an enneagon is shown slightly distorted in Fig. 2(b). The vertices, edges and face planes of the 3^9 remain unaltered, and as three such enneagons meet at each vertex, and each vertex has as its vertex figure an equilateral triangle, it follows that the polyhedron having the enneagons as its faces satisfies the definition of vertex-regularity. Hence there is a vertex-regular polyhedron 9^3, which occupies the same space as 3^9, and is in fact 3^9 thought of in another way.

The vertices of a 9^3 are also the centres of its faces, and again the vertices of a 3^9, so that the polyhedra 3^9 and 9^3 are reciprocal in the sense that the vertices of a 3^9 are the centres of the faces of a 9^3. The main point of interest is that the polyhedron 3^9 reciprocates in this sense into a polyhedron occupying the same space.

Balliol College, Oxford

[3]W. W. R. Ball, *Mathematical Recreations and Essays* (11th ed.), London, 1949, p. 147.

ON CERTAIN INTERSECTION PROPERTIES OF CONVEX SETS

V. L. KLEE, JR.

1. Introduction. A collection of $n+1$ convex subsets of a Euclidean space E will be called an *n-set in E* provided each n of the sets have a common interior point although the intersection of all $n+1$ interiors is empty. It is well-known that if $\{C_0, C_1\}$ is a 1-set, then C_0 and C_1 can be separated by a hyperplane. In the present note this result is generalized (Theorem I) by showing that if $\{C_0, \ldots, C_n\}$ is an n-set in E, then there is a variety V of deficiency n in E such that V intersects no set[1] Int C_i although in each direction away from V, V has a translate which intersects some set Int C_i. This theorem is then used to prove the converse (Theorem II) of Horn's recent generalization [2] of Helly's theorem [1] on the intersection of convex sets. The method of proof is essentially an elaboration of that of [3] and [4]. All theorems are stated only for a finite-dimensional Euclidean space E, although most of the proofs apply in rather general linear spaces.

2. A preliminary result. The following result will be useful in the sequel.

(2.1) *Suppose that $C_0, \ldots, C_n$ are closed convex subsets of E, each n of which have a point in common, and that $\bigcup_0^n C_i$ is convex. Then there is a point in common to all the C_i's.*

Proof. We may assume without loss of generality that all the C_i's are compact. For $n = 0$ the theorem is trivial. Now suppose it holds for $n=k-1$ and consider the case $n = k$. If $\bigcap_0^k C_i = \Lambda$ then C_0 and $P = \bigcap_1^k C_i$ are disjoint compact convex sets, so they can be separated by a hyperplane H disjoint from both of them. Let $C'_i = C_i \cap H (1 \leqslant i \leqslant k)$. For an arbitrary integer j between 1 and k let $X = \bigcap C_i \, (1 \leqslant i \leqslant k, i \neq j)$. Since each k of the C_i's have a point in common, X intersects C_0. And since furthermore $P \subset X$, X must intersect H and hence $\bigcap C'_i \neq \Lambda$ $(1 \leqslant i \leqslant k, i \neq j)$. But $\bigcup_1^k C'_i$ is convex, so it follows from the inductive hypothesis that $\bigcap_1^k C'_i \neq \Lambda$. Since this contradicts the fact that $P \cap H = \Lambda$, the proof is complete.

This remark may also be of interest.

(2.2) *Suppose that Γ is a collection of closed convex subsets of E such that*
 (i) *either Γ is finite or some set in Γ is compact;*
 (ii) *every finite subcollection of Γ has either a convex union or a non-empty intersection.*

Then there is a point in common to all sets of Γ.

Received April 4, 1950.

[1] Int X means the interior of X.

Proof. We need merely show that Γ has the finite intersection property, and this follows easily by an inductive argument which uses (2.1).

3. The "separation theorem" for n-sets. Let $x_0, x_1, \ldots, x_n$ be points of E; then $[x_0, x_1, \ldots, x_n]$ will denote the convex hull of the set $\{x_0, x_1, \ldots, x_n\}$, and $[x_0, x_1, \ldots, x_n] \equiv [x_0, x_1, \ldots, x_n] - \{x_0\}$, etc. (The sign $-$ is used for both set and vector differences, since in each case the meaning is clear from the context; $+$ is used for vector sum; $\cup$ for set union.)

The proof of Theorem I is effected by means of two lemmas, the first of which is the following:

(3.1) *If $\{C_0, \ldots, C_n\}$ is an n-set in E, then there are convex sets $K_i \supset C_i$ such that $\{K_0, \ldots, K_n\}$ is an n-set which covers E.*

Proof. We will show that if x is an arbitrary point of E then there are convex sets $C'_i \supset C_i$ such that $\{C'_0, \ldots, C'_i\}$ is an n-set and, in addition, $x \in \cup_0^n C'_i$; (3.1) follows from this fact by a straightforward application of Zorn's lemma.

For $0 \leqslant j \leqslant n$, $D_j \equiv \cap_{i \neq j} \text{Int } C_i$. If for some j we cannot have $d_j \in (c_j, x)$, with $d_j \in D_j$ and $c_j \in C_j$, then we merely let C'_j be the convex hull of $C_j \cup \{x\}$, $C'_i = C_i$ for $i \neq j$, and the sets C'_i will have the desired properties. Suppose this is not the case; that is, that there are points $d_0, \ldots, d_n, c_0, \ldots, c_n$ such that for $0 \leqslant i \leqslant n$, $c_i \in C_1$, and $d_i \in D_i \cap (c_i, x)$. For each j let $X_j = (c_j, d_0, \ldots, d_{j-1}, d_{j+1}, \ldots, d_n)$. Then $X_j \subset \text{Int } C_j$. But by use of Cramer's rule it can be shown that all the X_i's have a point in common and hence that $\cap_0^n \text{Int } C_i \neq \Lambda$, which is a contradiction, completing the proof of (3.1).

A linear subset of E is called a *subspace*, and each translate of a subspace is a *variety*. The *deficiency* in E of a subspace (and of its translates) is the dimension of a subspace complementary to it.

(3.2) *Suppose that $\{K_0, \ldots, K_n\}$ is an n-set which covers E and that $V \equiv \cap_0^n \overline{K}_i$. Then $V = E - \cup_0^n \text{Int } K_i$, and is a variety of deficiency n in E.*

Proof. Let $W = E - \cup_0^n \text{ Int } K_i$ and (for each j) $\pi_j = \cap_{i \neq j} \text{ Int } K_i$. From (2.1) we see that V is non-empty. We show first that $V \subset W$. For if not, there is a point p and an integer j such that $p \in V \cap \text{Int } K_j$. Let $q \in \pi_j$. Then, since for each $i \neq j$ we have $p \in \overline{K}_i$ and $q \in \text{Int } K_i$, $(p, q) \subset \pi_j$. But also $p \in \text{Int } K_j$, so (p, q) intersects $\text{Int } K_j$ and $\cap_0^n \text{ Int } K_i \neq \Lambda$, which is a contradiction.

To see that $W \subset V$, let $y \in W$ and $z \in \pi_j$ for some j. Consider an arbitrary point x such that $y \in (x, z)$. If, for any $i \neq j$, $x \in K_i$, then we have $y \in \text{Int } K_i$, which contradicts the fact that $y \in W$. Hence $x \in K_j$. Thus we have shown that $y \in \overline{K}_j$ for each j, and consequently $y \in V$. Since $W \subset V$ and $V \subset W$, $V = W$.

Obviously V is convex. To prove that it is actually a variety we must show that if $y \in V$, $z \in V$, and $y \in (x, z)$, then $x \in V$. But if x non $\in V$,

then (since $V = W$) $x \in \text{Int } K_j$ for some j and hence $y \in \text{Int } K_j$, which contradicts the fact that $y \in W$. Hence V is a variety and it remains only to show that the deficiency of V in E is n.

We assume without loss of generality that V contains the origin. Let S be a subspace of E complementary to V. Each point x of E has a unique expression in the form $v_x + x^*$ where $v_x \in V$ and $x^* \in S$. From the fact that $V = W$ it follows that no translate of V other than V itself can intersect all the sets K_i. This in turn implies that $\{K^*_0, K^*_1, \ldots, K^*_n\}$ is an n-set in S. Now Helly's theorem applies to an arbitrary *finite* collection of convex sets even though they may not be compact, so we can conclude that S is at least n-dimensional.

If $p_i \in \pi_i$ for each i then the variety U determined by $\{p_0, \ldots, p_n\}$ is n-dimensional. Let x be an arbitrary point of E. For a sufficiently small positive t we have $p_i + tx \in \pi_i$ for each i. Now for each i let $K_i{}^t = K_i \cap (U + tx)$. $\{K_0{}^t, K_1{}^t, \ldots, K_n{}^t\}$ is an n-set in $U + tx$, so V must intersect $U + tx$. From this it follows that V must intersect every translate of U, and hence that the deficiency of V in E is no greater than n. This completes the proof of (3.2).

THEOREM I. *If $\{C_0, \ldots, C_n\}$ is an n-set in E, then there is a variety V of deficiency n in E such that*

(a) *V intersects no set* $\text{Int } C_i$;

(b) *if V' is any variety of deficiency $n-1$ which contains V, and H is either of the half-spaces into which V separates V', then H intersects some set* $\text{Int } C_i$.

Proof. Let the K_i's be as in (3.1) and V as in (3.2). For each j let $z_j \in D_j = \bigcap_{i \neq j} C_i$ and let S be the variety determined by $\{z_0, \ldots, z_n\}$. S is a variety which is intersected by V in a single point P, and $\sigma = [z_0, \ldots, z_n]$ is an n-simplex whose boundary (relative to S) is contained in the union of the C_i's. In fact, if F_j is the face determined by $\{z_i | i \neq j\}$, then $F_j \subset \text{Int } C_j$. Now if V' is a variety of deficiency $n - 1$ (in E) which contains V, then V' intersects S in a line through P. Hence to prove Theorem I we need merely show that $P \in \sigma$. But if P non $\in \sigma$ then for some j there is a point $q \in F_j$ such that either $q \in (P, z_j)$ or $z_j \in (q, P)$. In the first case this implies that $P \in \pi_j$, in the second that $z_j \in \text{Int } C_j$, so in either case that $\bigcap_0^n \text{Int } C_i \neq \Lambda$, which is a contradiction, completing the proof.

4. Horn's generalization of Helly's theorem.

THEOREM II. *Suppose that Γ is a collection of compact convex subsets of E. Then the following statements are equivalent*:

(i) *every n members of Γ have a point in common*;

(ii) *each variety of deficiency n in E is contained in a variety of deficiency $n-1$ which intersects every member of Γ*;

(iii) *each variety of deficiency $n-1$ has a translate which intersects every member of Γ.*

Proof. That (i) implies (ii) is Horn's Theorem 4 [2; p. 928]. This (and that (i) implies (iii)) can also be proved by an argument which, like the proof of (2.1) above, leans heavily on the separation theorem and closely resembles Helly's proof [1] of his theorem.

To see that (ii) implies (iii), consider an arbitrary variety V of deficiency $n-1$ in E. We assume without loss of generality that V contains the origin. Now let S be a subspace of V, of deficiency 1 in V, and let $p \in V-S$. It follows from (ii) that for each integer k there is a variety V_k of deficiency $n-1$ in E such that $S + kp \subset V_k$ and V_k intersects all the members of Γ. Now the sequence of sets $V_1, V_2, \ldots$, must have a convergent subsequence [5, pp. 10-12]; say $\lim_{i\to\infty} V_{n_i} = W$. Since the members of Γ are compact, W intersects each member of Γ, and it follows by a simple argument that W is a translate of V. Thus (ii) implies (iii).

To complete the proof of Theorem II we show that if (i) is false, then so is (iii). For if (i) is false then for some $m < n$ there are sets $B_0, \ldots, B_m$ in Γ and open convex sets $C_i \supset B_i$ such that $\{C_0, \ldots, C_m\}$ is an m-set. Let V be the variety of Theorem I. Then V is of deficiency $m \leqslant n$ and intersects no set Int C_i. We assume without loss of generality that V contains the origin. Now consider an arbitrary translate $V + x$ of V (for x non $\in V$). For some $t > 0$, $V - tx$ intersects some set Int C_j. But then if $V + x$ intersects B_j, V must intersect Int C_j, which is a contradiction. Hence $V + x$ does not intersect all the sets B_i, and the proof is complete.

References

[1] E. Helly, *Über Mengen konvexer Körper mit gemeinschaftlichen Punkten*, Jber. Deutschen Math. Verein., vol. 32 (1923), 175-176.

[2] Alfred Horn, *Some generalizations of Helly's theorem on convex sets*, Bull. Amer. Math. Soc., vol. 55 (1949), 923-929.

[3] V. L. Klee, Jr., *Convex sets in linear spaces*, Dissertation (University of Virginia, 1949).

[4] Shizuo Kakutani, *Ein Beweis des Satzes von M. Eidelheit über konvexe Mengen*, Proc. Imp. Acad., Tokyo, vol. 13 (1937), 93-94.

[5] G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloquium Publications, vol. 28, (New York, 1942).

University of Virginia

ON THE TOPOLOGICAL THEORY OF FUNCTIONS

JAMES A. JENKINS

1. Introduction. The present paper constitutes a continuation of the ideas and methods of M. Morse and M. Heins [1]. As in that work the subject treated is the theory of deformation classes of meromorphic functions and interior transformations. There the functions considered were defined over the open disc $|z| < 1$ and had only a finite number of zeros, poles and branch point antecedents. It is possible to transfer the results obtained to the situation where the domain of definition is any simply-connected domain of hyperbolic type or, alternatively, of parabolic type. We shall be concerned principally with restricted deformations of functions having these same domains of definition but which are allowed to possess infinitely many zeros, poles and branch point antecedents. The same invariants as in [1] serve to characterize the restricted deformation classes.

The proof of this result presents essential difficulties not met with in the case of finitely many zeros, poles and branch point antecedents. It is no longer possible to use the Lagrange interpolation formula to construct canonical meromorphic functions. In order to prove the existence of a deformation for the interior transformations to be treated it is necessary to give a uniform version of Mittag-Leffler's theorem. In this same connection, the ordinary form of the Tietze deformation theorem is inadequate and it is necessary to give a uniform version of it also.

I wish to thank Professor Morse for his helpful comments.

2. Fundamental definitions. We shall begin by recalling certain fundamental definitions and results.

An *interior transformation* $w = f(z)$ defined on an open set S of the complex z-sphere is a continuous map of S into the complex w-sphere such that given any point z_0 of S there exists a sense-preserving homeomorphism $\phi(z)$ from a neighbourhood N of z_0 to another neighbourhood N_1 of z_0 with z_0 fixed such that the function $f[\phi(z)] = F(z)$ is analytic on N except for a possible pole at z_0 and is not identically constant. The transformation f is said to have a *zero* or *pole* at z_0 according as F has a zero or pole at z_0 and the *order* in each case is taken to be the corresponding order for F. If z_0 is the antecedent of a branch point of the rth order of the inverse of F, z_0 is said to be an *antecedent of a branch point of the rth order* of the inverse of f. These definitions extend at once to a Riemann domain by applying them in the plane of a local uniformizing parameter. We shall be concerned with the case that S is a simply-connected domain, of hyperbolic or parabolic type, and shall consider mero-

[1]Received April 3, 1950, while the author was a Frank B. Jewett Fellow at the Institute for Advanced Study.

morphic functions and interior transformations which may have infinitely many zeros, poles and branch point antecedents. The totality of such points will be called the *characteristic set* (a) of the function. We shall assume that all zeros, poles and branch points involved are of order 1.

By a *deformation* of an interior transformation f we shall mean a family of interior transformations

$$w = F(z, t) \qquad (z \in S,\ 0 \leqslant t \leqslant 1)$$

depending continuously on the point z in S and the deformation parameter t together and having

$$F(z, 0) = f(z) \qquad (z \in S).$$

We shall confine our attention to ***restricted*** deformations, namely those for which $F(z, t)$ has for every t the same zeros, poles and branch point antecedents as $f(z)$. Interior transformations which admit a (restricted) deformation one into the other will be said to belong to the same (restricted) deformation class.

The fundamental concept used in defining the invariants employed to characterize the deformation properties of interior transformations is that of the difference order $d(k)$ of a locally simple arc k. We consider arcs k which are represented by continuous and locally (1, 1) images

$$w(t) = u(t) + iv(t) \qquad (0 \leqslant t \leqslant 1)$$

of the interval [0, 1] and which intersect their end points $w(0) = a$ and $w(1) = b$ only when $t = 0$ and $t = 1$ respectively. We call such arc *locally simple* provided there exists a constant $e > 0$ such that any subarc of k whose diameter is less than e is simple. Such a constant e is called a *norm of local simplicity* of k. A set of locally simple curves which admit the same norm of local simplicity will be termed *uniformly locally simple.*

An *admissible deformation* of a locally simple arc k is defined by a family of arcs represented in the form

$$w(t) = H(t, \lambda) \qquad (0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1)$$

where H depends continuously on t and λ together and satisfies

$$H(0, \lambda) = a,\ H(1, \lambda) = b \qquad (0 \leqslant \lambda \leqslant 1).$$

The arc associated with the parameter λ will be called k^λ. Then the arcs k^λ are to be uniformly locally simple and are to intersect a and b as end points only.

Two arcs connected by such a deformation are said to belong to the same *deformation class.* This property is independent of the particular parametric representation of the arcs [2, Lemma 28.1].

We proceed to the definition of the *difference order* of a locally simple arc joining two finite points a and b $(a \neq b)$.

A subarc of k will be defined by an interval (σ, τ) for t. If this subarc is

simple the vector $w(\tau) - w(\sigma)$ has a well defined direction and this direction will vary continuously with σ and τ so long as the arc given by (σ, τ) remains simple and $\sigma < \tau$. Such a variation is termed an admissible chord variation and if we suppose that the angle

$$\frac{1}{2\pi} \arg [w(\tau) - w(\sigma)] \tag{1}$$

has been chosen so as to change continuously in the course of the variation, then the algebraic increment of the angle depends only on the initial and final simple subarcs.

Let k_a and k_b be respectively proper simple subarcs of k of which the initial point of k_a is a and the terminal point of k_b is b. Let the chord subtending k_a vary admissibly into the chord subtending k_b. Let $P(k, k_a, k_b)$ represent the accompanying algebraic increment of (1). Let the algebraic increment of

$$\frac{1}{2\pi} \arg [w(t) - a] \tag{2}$$

as t increases from its terminal value t_a on k_a to 1 be denoted by $Q_a(k, k_a)$ and let the algebraic increment of

$$\frac{1}{2\pi} \arg [w(t) - b] \tag{3}$$

as t increases from 0 to its initial value t_b on k_b be denoted by $Q_b(k, k_b)$. Then the difference order $d(k)$ of k is defined by

$$d(k) = P(k, k_a, k_b) - Q_a(k, k_a) - Q_b(k, k_b). \tag{4}$$

The difference order $d(k)$ as defined by (4) where $a \neq b$ is an integer which is independent of the choice of k_a and k_b among proper simple subarcs of k with end points as prescribed. It is further independent of any admissible deformation of k [1, Lemma 3.3].

The case a finite, $b = \infty$. In this case continuity, local simplicity and related concepts must be interpreted appropriately in terms of the w-sphere. For a locally simple arc k joining the finite point a to the point at infinity, there will exist a value τ with $0 < \tau < 1$ such that the subarc corresponding to values of t with $\tau \leqslant t < 1$ is simple in the finite w-plane. On this arc $w(t)$ becomes infinite as t tends to 1. Let $k(s)$ denote the subarc of k defined by $(0, s)$ with $0 < s < 1$. Then $d(k(s))$ is well defined and is clearly independent of s provided $s > \tau$ so that we may set

$$d(k) = d(k(s)) \qquad (\tau < s < 1).$$

This quantity possesses invariance under admissible deformations as before.

The case $a = b$ (finite). Let $P(k, k_a, k_b)$ be as above and let $Q_a(k, k_a, k_b)$ equal the algebraic increment of

$$\frac{1}{2\pi} \arg [w(t) - a]$$

as t varies monotonically from its terminal value t_a on k_a to its initial value t_b on k_b. Then the difference order $d(k)$ of k is defined by

$$d(k) = P(k, k_a, k_b) - Q_a(k, k_a, k_b).$$

This quantity has the following properties. The value of $d(k)$ when $a = b$ is equal to $\frac{1}{2}$ modulo 1. It is independent of the choice of k_a and k_b among proper simple subarcs of k with end points as prescribed and is invariant under admissible deformations of k [1, Lemma 3.5].

3. The invariants. We shall for the present confine our attention to the special hyperbolic case where S is the open disc $|z| < 1$. We will consider an interior transformation $f(z)$ defined on S and possessing zeros and poles which we denote by a_j and branch point antecedents which we denote by b_j. In general the index will run, in the first case, through all non-negative integers and, in the second, through all positive integers. Exceptionally either sequence might terminate, but since this would result only in a considerable simplification of the situation we shall not mention this possibility explicitly. Evidently the points of the characteristic set can have no limit point in the interior of the unit circle. We shall assume that there is at least one zero which we take to be a_0. If there were poles but no zeros we would consider the reciprocal function. For each a_j with $j > 0$ we consider a simple curve h_j joining a_0 to a_j on S and meeting no other point of the characteristic set. Two such curves h_j will be said to be of the same topological type if they can be deformed isotopically one into the other on S without intersecting the characteristic set of f elsewhere than in the end points of h_j. Let h_j^f denote the image of h_j under f. The curve h_j^f will join $f(a_0)$ to $f(a_j)$ in the w-plane and be locally simple. An isotopic deformation of h_j through curves of the same topological type will induce an admissible deformation in the curves h_j^f. Thus the difference order $d(h_j^f)$ is invariant under such deformations. It will, however, in general be different for curves of different topological types.

By a slight generalization of the Weierstrass product representation for integral functions as it is usually given we can construct one function regular in S which has simple zeros at the points a_j and no other zeros in S and a second function regular in S which has simple zeros at the points b_j and no other zeros in S. There is a good deal of arbitrariness in the choice of these functions and we shall later want to make our choice subject to certain conditions. However we shall always use two definite functions $A(z,a)$ and $B(z,a)$, say, kept fixed for all time for a given characteristic set (a). We shall not go further into the above generalization of the product representation at this point since we shall later have to prove a stronger result which includes this.

Now let us set

$$C_j(z, \alpha) = \frac{(z - a_0)(z - a_j)B(z, \alpha)}{A(z, \alpha)} \qquad (j > 0)$$

where $C_j(z, \alpha)$ is assumed to take its limit values at a_0 and a_j, namely

$$C_j(a_0, \alpha) = \frac{(a_0 - a_j)B(a_0, \alpha)}{A'(a_0, \alpha)}, \; C_j(a_j, \alpha) = \frac{(a_j - a_0)B(a_j, \alpha)}{A'(a_j, \alpha)} \qquad (j > 0).$$

Corresponding to a variation of z along h_j we set

$$V(h_j) = \frac{1}{2\pi} [\arg C_j(z, \alpha)]_{z=a_0}^{z=a_j}.$$

Regardless of the arguments used

$$V(h_j) \equiv \frac{1}{2\pi} [\arg C_j(a_j, \alpha) - \arg C_j(a_0, \alpha)] \qquad (\text{mod } 1).$$

In [1] there is proved the following theorem which we quote: [1, Theorem 5.2].

THEOREM A. *The value of the difference*

$$d(h_j f) - V(h_i)$$

is independent of h_j among simple curves which join a_0 to a_j without intersecting the other points of the characteristic set (α).

The proof of this theorem in no wise uses the assumption there in force that there are only finitely many points in the characteristic set but merely that between two choices for h_j lie only finitely many points of this set. This condition is certainly satisfied here and in all cases with which we shall deal. This enables us to make the following definition.

We set

$$J_j(f, \alpha) = d(h_j f) - V(h_j) \qquad (j > 0)$$

for any simple arc h_j which joins a_0 to a_j without intersecting $(\alpha) - (a_0, a_j)$. The numbers J_j are independent of the choice of h_j among admissible arcs h_j and of restricted deformations of f. We should remark that these quantities can be defined analogously for S the finite z-plane and their definition extends at once to any domain conformally equivalent to either of the preceding.

Our objective is to characterize the deformation properties of meromorphic functions and interior transformations in terms of the numbers J_j which we call the *invariants* of the function in question. This we will do by means of the following theorem.

THEOREM B. *A necessary and sufficient condition that the meromorphic functions or interior transformations f_1 and f_2, defined on S and possessing the same characteristic set, belong to the same restricted deformation class is that*

$$J_j(f_1, \alpha) = J_j(f_2, \alpha) \qquad (j > 0).$$

In the case of meromorphic functions the deformation can be carried out through meromorphic functions alone.

The necessity is obvious from what we have observed above. We shall next treat the sufficiency in the case of meromorphic functions.

4. Meromorphic functions. Let us suppose that the function f is meromorphic on S and possesses the characteristic set (α). Then the function ϕ defined by the equation

$$\frac{f'(z)}{f(z)} = \phi(z)\frac{B(z, \alpha)}{A(z, \alpha)} \tag{5}$$

is regular on S apart from removable singularities and never zero. We term ϕ the *residual function* of f.

We have immediately at our disposal the following theorem [1, Theorem 10.1].

THEOREM C. *The algebraic increment of the argument of the residual function ϕ of f as z traverses a simple regular arc h_i leading from a_0 to a_j on S is equal to $2\pi J_j(f, \alpha)$.*

On multiplying the two sides of (5) by $z - a_j$ and letting z tend to a_j as limit, we find that

$$\phi(a_j) = e_j \frac{A'(a_j, \alpha)}{B(a_j, \alpha)} \qquad (j \geqslant 0) \tag{6}$$

where $e_j = 1$ if a_j is a zero and $e_j = -1$ if a_j is a pole. Regarding this latter quantity as a function determined by the characteristic set we define

$$e_j \frac{A'(a_j, \alpha)}{B(a_j, \alpha)} = g_j(\alpha) \qquad (j \geqslant 0).$$

The following lemma is proved precisely as in [1] provided we include as a condition of admissibility that the points of the characteristic set have no limit point in the interior of S [1, Lemma 10.2].

LEMMA 1. *Corresponding to an arbitrary admissible characteristic set* (α) *and to a function ψ which is non-null and regular on S and satisfies* (6) *in terms of* (α), *there exists a function F which is meromorphic on S, possesses the characteristic set* (α) *and for which the residual function is ψ.*

We will employ the convention of denoting, for a complex quantity X, by Arg X that value of the argument for which

$$0 \leqslant \operatorname{Arg} X < 2\pi$$

and similarly

$$\operatorname{Log} X = \log |X| + i \operatorname{Arg} X.$$

Now we may express the invariants of any interior transformation by

$$J_j(f, \alpha) = \frac{1}{2\pi}\{\operatorname{Arg} g_j(\alpha) - \operatorname{Arg} g_0(\alpha)\} + I_j(f, \alpha)$$

where the numbers $I_j(f, \alpha)$ are integers invariant under restricted deformations of f and uniquely determined by f and (α). This is a consequence of the facts that the first term differs from $-V(h_j)$ by half of an odd integer or by an integer according as a_j is a zero or a pole and that in these same cases $d(h_j f)$ is respectively half of an odd integer or an integer.

The following theorem shows that all possible values of the integers I_j are realized for meromorphic functions and hence more generally for interior transformations.

THEOREM D. *Corresponding to any admissible characteristic set* (α) *and arbitrary sequence* $\{r_j\}$ *of integers, there exists a function* $F(z, \alpha, r)$ *which is meromorphic in* z *on* S, *whose characteristic set is* (α) *and whose invariants* $I_j(F, \alpha) = r_j$ $(j > 0)$.

Indeed, let us set

$$c_j = c_j(\alpha, r) = \operatorname{Log} g_j(\alpha) + 2\pi r_j i \qquad (j \geqslant 0)$$

where we take $r_0 = 0$. Let us further set

$$A'_j = A'(a_j, \alpha).$$

Since all zeros of $A(z, \alpha)$ are simple, none of these quantities are zero. Thus we can construct a function $Q(z, \alpha, c)$ which has simple poles at the points $a_j (j \geqslant 0)$, the corresponding residues being c_j/A'_j and which is regular elsewhere in S. This is done by a slight variation of the proof of Mittag-Leffler's theorem as it is usually given [6]. Since, in any case, we shall later need a much stronger result which will imply this fact, we shall not go into this variation here.

Now the function

$$P(z) = P(z, \alpha, c) = A(z, \alpha)\, Q(z, \alpha, c)$$

is regular for $z \in S$ and has the values

$$P(a_j) = c_j \qquad (j \geqslant 0).$$

The function

$$\psi(z) = \exp P(z)$$

gives the residual function corresponding to a function $F(z, \alpha, r)$ of the desired type. Indeed,

$$\psi(a_j) = \exp P(a_j) = \exp c_j = g_j(\alpha) = e_j \frac{A'(a_j, \alpha)}{B(a_j, \alpha)} \qquad (j \geqslant 0),$$

and $\psi(z)$ is regular and non-zero on S. Hence it is a residual function by Lemma 1. Let $F(z, \mathfrak{a}, r)$ be the corresponding meromorphic function. Since ψ is non-zero on S there exists a single-valued continuous branch of arg ψ over S and, by Theorem C, for any such

$$2\pi J_j(\mathrm{F}, \mathfrak{a}) = \arg \psi(a_j) - \arg \psi(a_0).$$

Then by the definition of $\psi(z)$

$$\begin{aligned} 2\pi J_j(F, \mathfrak{a}) &= \mathfrak{I}(P(a_j)) - \mathfrak{I}(P(a_0)) \\ &= \mathfrak{I}(c_j) - \mathfrak{I}(c_0) \\ &= (\operatorname{Arg} g_j(\mathfrak{a}) + 2\pi r_j) - \operatorname{Arg} g_0(\mathfrak{a}). \end{aligned}$$

Hence

$$I_j(F, \mathfrak{a}) = r_j \qquad (j > 0)$$

as required.

There is a certain arbitrary character in the definition of $F(z, \mathfrak{a}, r)$ since in the definition of $A(z, \mathfrak{a})$, $B(z, \mathfrak{a})$ and $Q(z, \mathfrak{a}, c)$ were involved certain arbitrary choices. However, to a certain extent $F(z, \mathfrak{a}, r)$ can be used as a canonical function associated with the chosen sets $(\mathfrak{a})$ and (r), since, when these are given, we may make a perfectly definite choice of the functions A, B and Q.

We shall now give the proof of the part of Theorem B which relates to meromorphic functions. It is the same as the corresponding proof in [1] but we include it for the sake of completeness.

Let, then, f_1 and f_2 be two functions meromorphic on S with the same characteristic set $(\mathfrak{a})$ for which

$$J_j(f_1, \mathfrak{a}) = J_j(f_2, \mathfrak{a}) \qquad (j > 0).$$

Let ϕ_1 and ϕ_2 be the corresponding residual functions and let us set

$$\psi(z, t) = \exp\{(1 - t) \log \phi_1(z) + t \log \phi_2(z)\} \qquad (0 \leqslant t \leqslant 1)$$

where $\log \phi_1$ and $\log \phi_2$ are continuous branches with

$$\log \phi_1(a_0) = \log \phi_2(a_0). \tag{7}$$

The latter can be arranged since

$$\phi_1(a_0) = \phi_2(a_0).$$

Clearly $\psi(z, t)$ is non-zero and regular on S. Equality of the invariants for f_1 and f_2 implies that

$$\arg \phi_1(z) \Big|_{a_0}^{a_j} = \arg \phi_2(z) \Big|_{a_0}^{a_j} \qquad (j > 0)$$

or again

$$\log \phi_1(z) \Big|_{a_0}^{a_j} = \log \phi_2(z) \Big|_{a_0}^{a_j} \qquad (j > 0).$$

Then from (7) we deduce

$$\log \phi_1(a_j) = \log \phi_2(a_j) \qquad (j > 0).$$

Thus

$$\psi(a_j, t) = \exp \log \phi_1(a_j) = \phi_1(a_j) \qquad (j > 0).$$

By Lemma 1 this implies that $\psi(z, t)$ is the residual function of a meromorphic function with characteristic set (α)

$$f(z, t) = \exp\left(\int \psi(z, t) \frac{B(z, \alpha)}{A(z, \alpha)} dz\right) \qquad (0 \leqslant t \leqslant 1).$$

This means that

$$f(z, 0) = C_1 f_1(z) \qquad (C_1 \neq 0),$$

$$f(z, 1) = C_2 f_2(z) \qquad (C_2 \neq 0),$$

with C_1, C_2 constants. The function

$$\frac{f(z, t)}{C_1^{1-t} C_2^{t}} \qquad (0 \leqslant t \leqslant 1)$$

gives the required meromorphic deformation of f_1 into f_2.

5. Interior transformations. We now proceed to the classification of interior transformations under restricted deformation. It is here that the infinitude of the characteristic set first makes itself felt in an essential manner since there are no longer canonical meromorphic functions depending in a trivial manner on the characteristic set.

We make here a remark which will be useful later also: if we can prove the classification theorem after a preliminary homeomorphism of S it follows immediately for the original situation.

Indeed, equality of the invariants of two interior transformations with the same characteristic set is preserved by such an operation since in their definition, if we use arcs corresponding under the homeomorphism for the evaluation, the term $d(h_j^f)$ is unaltered and the term $-V(h_j)$ is changed by an amount independent of the particular interior transformation. Further, if we can perform the deformation after the homeomorphism we can transform it back to the original situation and conversely.

Now we may assume that all points of the characteristic set (α) lie in the semi-circle where $\Re z \geqslant 0$. Indeed, if we draw every circle with centre $z = 0$ and passing through a point of the characteristic set, then on these circles we can deform the identity mapping isotopically into a homeomorphism of the circle for which the characteristic points have their images in $\Re z \geqslant 0$ and then extend this homeomorphism over the intervening circular rings [5].

Now we can prove:

Lemma 2. *Any interior transformation f^* of S admits a restricted deformation into a function f for which a sense-preserving homeomorphism ζ of S exists such that $f\zeta$ is meromorphic on S.*

Indeed, by virtue of what we may call the "uniformization theorem" [4] there exists a homeomorphism σ from a simply-connected domain R of the complex plane to S such that $f^*\sigma$ is meromorphic on R. If R is of hyperbolic type it can be mapped conformally onto S and σ is carried into a homeomorphism ζ as desired. If R is of parabolic type we can perform a preliminary restricted deformation of f^* and reduce this case to the preceding one. Indeed let θ^t be an isotopic deformation of S onto $S' = \{z \mid |z| < 1, \Re z > -1 + \delta, 0 < \delta < \frac{1}{2}\}$, leaving fixed the points of S with $\Re z > -1 + 2\delta$. Then θ^1 will be a homeomorphism from S to S' and the composite function $f^*\theta^t$ constitutes a restricted deformation of f^*. The Riemann image under this new function, as a proper subdomain of the Riemann image under f^*, is of hyperbolic type and thus comes under the case first treated.

It is at this stage enough to show that an interior transformation of the type of the function f of the lemma admits a restricted deformation into a meromorphic function.

Let f be an interior transformation for which there exists a sense-preserving homeomorphism ζ of S onto itself such that $f\zeta$ is meromorphic on S. Let η be the inverse homeomorphism of ζ on S. We will denote by $C(r)$ the circle $|z| = r$ and by $L(r)$ its image under η. Further $m(r)$ will denote the minimum of $|z|$ on $L(r)$. Since, as $r \to 1$, $L(r)$ passes outside of every compact subset of S, we have

$$\lim_{r\to 1} m(r) = 1.$$

We will need the following lemma, the present simple proof of which is due to a suggestion by Professor Morse.

Lemma 3. *There exists an isotopic deformation η^t $(0 \leqslant t \leqslant 1)$ of S which deforms the identity mapping of S into the homeomorphism η and which satisfies the following condition: if $L^t(r)$ and $m^t(r)$ are the quantities for η^t analogous to $L(r)$ and $m(r)$ for η then*

$$\lim_{r\to 1} m^t(r) = 1$$

uniformly in t.

Indeed let us consider the space $\bar{S}$ obtained from S by Alexandroff's compactification. That is, we adjoin to S a point P and define a neighbourhood of P to consist of P together with the complement of a compact set in S. The space $\bar{S}$ is topologically equivalent to a 2-sphere and the homeomorphism η extends at once to a sense-preserving homeomorphism $\bar{\eta}$ of $\bar{S}$ having P as fixed point. It is well known that $\bar{\eta}$ can be generated by an isotopic deformation from the identity $\bar{\eta}^t$ $(0 \leqslant t \leqslant 1)$ on $\bar{S}$ where P is fixed under each $\bar{\eta}^t$ [3]. Given any neighbourhood W of P, since $\bar{\eta}^t(z)$ is continuous in z and t together, for each t', $0 \leqslant t' \leqslant 1$, we can find a neighbourhood $U^{t'}$ of P and an interval $I^{t'}$ open with respect to $[0, 1]$ containing t' such that $\bar{\eta}^t(z)$ maps the direct product $U^{t'} \times I^{t'}$ into W. We can cover the interval $[0, 1]$ by a finite number

of such intervals and let $U_1, \ldots, U_k$ be the corresponding neighbourhoods of P. The intersection of these neighbourhoods is mapped into W by $\bar{\eta}^t$, regarded as a function of z only, for all t. It is clear that by restricting $\bar{\eta}^t$ to the complement of P, that is, passing back to S we obtain the desired isotopic deformation η^t.

The argument will now proceed in a number of steps. We will make use of our notations defined previously. For a deformation η^t of the type in Lemma 3 we will denote by a_n^t, b_n^t the respective images of a_n, b_n under η^t and by (a^t) the corresponding characteristic set.

(i) *The product*

$$\prod_{n=0}^{\infty}\left((1-\frac{z}{a_n{}^t}\right)\exp\left\{\sum_{j=1}^{n}\frac{1}{j}\left(\frac{z}{a_n{}^t}\right)^j\right\} = A(z, a^t) \tag{8}$$

is a continuous function of z and t together.

In order to prove continuity for $z = z_0$ it is enough to choose R with $|z_0| < R < 1$ and allow only variations of z within the circle $|z| < R$. The general factor in the product (8) can be written

$$\exp\left\{-\sum_{n+1}^{\infty}\frac{1}{j}\left(\frac{z}{a_n{}^t}\right)^j\right\} = \exp v_n(z),$$

say. By the condition on η^t we have $a_j{}^t = \theta_j{}^t a_j$ where $|\theta_j{}^t| \to 1$ uniformly in t as $j \to \infty$. Now

$$\begin{aligned}|v_n(z)| &\leqslant \sum_{n+1}^{\infty}\left|\frac{z}{a_n{}^t}\right|^j = \left|\frac{z}{a_n{}^t}\right|^{n+1}\Big/\left(1-\left|\frac{z}{a_n{}^t}\right|\right)\\ &\leqslant \theta\left|\frac{z}{a_n{}^t}\right|^{n+1}\end{aligned}$$

provided n is large enough, namely so that $|a_n{}^t| \geqslant R/(1-\epsilon)$ for all t, $1 > \epsilon > 0$, $\theta = \theta(\epsilon)$ independent of n. This shows that the infinite product (8) converges uniformly and absolutely in $|z| \leqslant R$, $0 \leqslant t \leqslant 1$. In particular, the product

$$\prod_{N+1}^{\infty}\left(1-\frac{z}{a_n{}^t}\right)\exp\left\{\sum_{j=1}^{n}\frac{1}{j}\left(\frac{z}{a_n{}^t}\right)^j\right\}$$

is arbitrarily close to 1 for N large enough, uniformly in $|z| \leqslant R, 0 \leqslant t \leqslant 1$. The terms

$$\prod_{n=0}^{N}\left(1-\frac{z}{a_n{}^t}\right)\exp\left\{\sum_{j=1}^{n}\frac{1}{j}\left(\frac{z}{a_n{}^t}\right)^j\right\}$$

depend continuously on z and t together for $|z| \leqslant R$, $0 \leqslant t \leqslant 1$, proving the result. We note that in case $a_n{}^t = 0$ for some t (which can occur only for finitely many n) the corresponding primary factor must be redefined in a suitable manner.

(ii) *The product*

$$\prod_{n=1}^{\infty}\left(1-\frac{z}{b_n{}^t}\right)\exp\left\{\sum_{i=1}^{n}\frac{1}{j}\left(\frac{z}{b_n{}^t}\right)^j\right\}=B(z,\mathfrak{a}^t)$$

is a continuous function of z and t together.

This is immediate by (i).

(iii) *The quantity $A'(a_j{}^t, \mathfrak{a}^t)$ is a continuous, non-zero function of t for each j ($j \geqslant 0$).*

Indeed

$$A'(z,\mathfrak{a}^t) = -\frac{1}{a_j{}^t}Q_j(z,t)+\left(1-\frac{z}{a_j{}^t}\right)Q'_j(z,t)$$

where

$$Q_j(z,t)=\exp\left\{\sum_{k=1}^{j}\frac{1}{k}\left(\frac{z}{a_j{}^t}\right)^k\right\}\prod_{n\neq j}\left(1-\frac{z}{a_n{}^t}\right)\exp\left\{\sum_{k=1}^{n}\frac{1}{k}\left(\frac{z}{a_n{}^t}\right)^k\right\}.$$

Further $Q_j(z, t)$ is a continuous function of z and t together and

$$A'(a_j{}^t,\mathfrak{a}^t) = -\frac{1}{a_j{}^t}\,Q(a_j{}^t,t)$$

which depends continuously on t. Once again if $a_j{}^t = 0$ for some t the corresponding primary factor must be suitably redefined.

(iv) *The quantity*

$$g_j(\mathfrak{a}^t) = e_j\,\frac{A'(a_j{}^t,\mathfrak{a}^t)}{B\,(a_j{}^t,\mathfrak{a}^t)}$$

is a continuous non-zero function of t for each j ($j \geqslant 0$).

Indeed $B(a_j{}^t, \mathfrak{a}^t) \neq 0$, $0 \leqslant t \leqslant 1$, and the result follows by (iii).

(v) Let us define

$$c_j{}^t = \log g_j(\mathfrak{a}^t) + 2\pi r_j i \qquad (j \geqslant 0),$$

where some definite determination of the logarithm is chosen at a fixed point for each j and the r_j are integers. The function

$$c_j{}^t/A'_j{}^t, \text{ with } A'_j{}^t = A'(a_j{}^t,\mathfrak{a}^t)$$

is a continuous function of t, $0 \leqslant t \leqslant 1$, and hence bounded in absolute value

$$|c_j{}^t/A'_j{}^t| \leqslant K_j \qquad (0 \leqslant t \leqslant 1).$$

There exists an increasing sequence of integers $\{j_n\}$ such that

$$\sum_{n=0}^{\infty}{}_n K\theta^{j_n}$$

converges for each θ, $0 \leqslant \theta < 1$. Indeed it is only necessary to arrange that $\lim_{n\to\infty} (K_n)^{1/j_n} = 1$. *Then the function, to play the role of $Q(z, \mathfrak{a}, c)$ in Theorem D,*

$$Q(z, \mathfrak{a}^t, c^t) = \sum_{n=0}^{\infty} \frac{c_n{}^t}{A'_n{}^t}\left\{\frac{1}{z - a_n{}^t} + \frac{1}{a_n{}^t}\left[1 + \frac{z}{a_n{}^t} + \ldots + \left(\frac{z}{a_n{}^t}\right)^{j_n-1}\right]\right\}$$

is a continuous function of z and t together (Q taken on the Riemann sphere). As before, to prove continuity of Q for a given value of z, say z_0, it is enough to choose R with $|z_0| < R < 1$ and allow only variations of z within the circle $|z| < R$. In this case

$$\begin{aligned}\left|\frac{1}{z - a_n{}^t} + \frac{1}{a_n{}^t}\left[1 + \frac{z}{a_n{}^t} + \ldots + \left(\frac{z}{a_n{}^t}\right)^{j_n-1}\right]\right| &\leqslant \left|\frac{z}{a_n{}^t}\right|^{j_n} \left|\frac{1}{a_n{}^t}\right| \frac{1}{1 - \left|\frac{z}{a_n{}^t}\right|} \\ &\leqslant \theta \left|\frac{z}{a_n{}^t}\right|^{j_n}\end{aligned}$$

provided n is large enough, namely so that $|a_n{}^t| \geqslant R/(1 - \epsilon)$ for all t, $1 > \epsilon > 0$, $\theta = \theta(\epsilon)$ independent of n. Thus the above series converges uniformly and absolutely for $|z| \leqslant R$, $0 \leqslant t \leqslant 1$. In particular, the sum

$$\sum_{n=N+1}^{\infty} \frac{c_n{}^t}{A'_n{}^t}\left\{\frac{1}{z - a_n{}^t} + \frac{1}{a_n{}^t}\left[1 + \frac{z}{a_n{}^t} + \ldots + \left(\frac{z}{a_n{}^t}\right)^{j_n-1}\right]\right\}$$

is arbitrarily small in absolute value uniformly in $|z| \leqslant R$, $0 \leqslant t \leqslant 1$, for N large enough. The terms

$$\sum_{n=0}^{N} \frac{c_n{}^t}{A'_n{}^t}\left\{\frac{1}{z - a_n{}^t} + \frac{1}{a_n{}^t}\left[1 + \frac{z}{a_n{}^t} + \ldots + \left(\frac{z}{a_n{}^t}\right)^{j_n-1}\right]\right\}$$

depend continuously on z and t together for $|z| \leqslant R$, $0 \leqslant t \leqslant 1$, proving the result.

Now let, as in the proof of Theorem D,

$$P(z, \mathfrak{a}^t, c^t) = A(z, \mathfrak{a}^t)\, Q(z, \mathfrak{a}^t, c^t)$$

and

$$\psi^t(z) = \exp\left\{P(z, \mathfrak{a}^t, c^t)\right\}.$$

Then

$$F^t(z) = \exp\left(\int \frac{\psi^t(z) B(z, \mathfrak{a}^t)}{A(z, \mathfrak{a}^t)}\, dz\right)$$

depends continuously on z and t together, $z \in S$, $0 \leqslant t \leqslant 1$. This is evident from the above considerations.

We are now in a position to prove the part of Theorem B which deals with interior transformations. Let f be an interior transformation on S, ζ a homeomorphism of S such that $f\zeta$ is meromorphic on S and η the inverse of ζ on S.

Let $\eta^t(0 \leqslant t \leqslant 1)$ be an isotopic deformation from the identity generating η and of the type of Lemma 3. We will denote the meromorphic function $f\zeta$ by λ: it has the characteristic set (a^1). The meromorphic function F^1 also has this characteristic set and the determination of the logarithms and the integers r_j in the definition of c_j^t can be chosen so that F^1 will have the same invariants as λ after the manner of Theorem D. From this point on we retain this choice of the determination. Then by the first part of the theorem there exists a restricted deformation of meromorphic type of λ into F^1 which we denote by λ^t $(0 \leqslant t \leqslant 1)$. The composite function

$$\lambda^t\eta \qquad (0 \leqslant t \leqslant 1)$$

will define a restricted deformation of $\lambda\eta \equiv f$ into $F^1\eta$. The function $F^t\eta^t$ always has the characteristic set (a) and as t decreases from 1 to 0 it defines a restricted deformation of $F^1\eta$ into F^0. This last is a meromorphic function. Combining these two deformations gives the required restricted deformation of f into a meromorphic function. By the way in which it was obtained, the latter naturally has the same invariants as f. This completes the proof of the theorem.

6. General domains. We will now consider how the results of the preceding three sections extend to domains other than the open disc. For this the remark made at the beginning of §5 is essential. Any simply-connected domain of hyperbolic type admits a conformal map on the unit disc and thus the results carry over as regards both meromorphic functions and interior transformations. In the finite z-plane the deformation properties of meromorphic functions are obtained simply by transferring word for word the proofs previously given. In the case of interior transformations we can return to the situation of the unit disc by a preliminary homeomorphism. As above the results extend at once to any simply-connected domain of parabolic type.

We may remark also that the extension to the case where zeros, poles and branch point antecedents of higher orders are allowed requires only formal modifications.

References

[1] Morse, M. and Heins, M., *Deformation classes of meromorphic functions and their extensions to interior transformations*, Acta Math., vol. 79 (1947), 51-103.

[2] Morse, M., *Topological methods in the theory of meromorphic functions*, Ann. of Math. Studies (Princeton, N.J., 1947).

[3] v. Kerékjártó, B., *Vorlesungen über Topologie* (Berlin, 1923).

[4] Stoïlow, S., *Leçons sur les principes topologiques de la théorie des fonctions analytiques* (Paris, 1938).

[5] Tietze, H., *Sur les représentations continues des surfaces sur elles-mêmes*, Comptes Rendus 157 (1913), 509-512.

[6] Dienes, P., *The Taylor Series* (Oxford, 1931).

Johns Hopkins University

A THEOREM ON DIVISION RINGS

IRVING KAPLANSKY

THE object of this note is to prove the following theorem.

THEOREM. *Let A be a division ring with centre Z, and suppose that for every x in A, some power (depending on x) is in Z: $x^{n(x)} \in Z$. Then A is commutative.*

This theorem contains as special cases three previously known results.

1. It includes Wedderburn's theorem that any finite division ring is commutative, and the generalization by Jacobson [3, Theorem 8] asserting that any algebraic division algebra over a finite field is commutative; for in such an algebra every non-zero element has some power equal to 1.

2. It includes a theorem of Emmy Noether, as generalized by Jacobson [3, Lemma 2], stating that any non-commutative algebraic division algebra contains an element separable over the centre; for otherwise a suitable p^mth power of every element would lie in the centre.

3. Hua [1, Theorem 7] has proved the special case of the theorem where the power n is independent of x, and the characteristic is at least n.

Although our theorem generalizes the two cited theorems of Jacobson, we are not giving a new proof of these theorems. In fact, we shall prove a preliminary lemma on fields which reduces the problem precisely to these two theorems.

LEMMA. *Let K be a field and L an extension of K, $L \neq K$, with the property that for every x in L, some power (the power depending on x) lies in K. Then L has prime characteristic, and it is either purely inseparable over K, or algebraic over its prime subfield.*

Proof. If L is indeed purely inseparable over K, there is of course nothing to prove. So suppose L contains an element y, y non $\in K$, which is separable over K. By a suitable isomorphism leaving K elementwise fixed, y can be sent into an element $z \neq y$ (of course z need not be in L). We have, say, $y^r \in K$ and and so $z^r = y^r$, whence $z = \epsilon y$ with $\epsilon^r = 1$. Suppose $(1 + y)^s \in K$; then similarly $1 + z = \eta(1 + y)$ with $\eta^s = 1$. We cannot have $\epsilon = \eta$, for then $\epsilon = 1$, $z = y$. So we may solve for y:

$$y = (1 - \eta)(\eta - \epsilon)^{-1}. \tag{1}$$

We see that y is algebraic over the prime subfield P of K. If k is any element of K, we can repeat this argument with $k + y$ instead of y, and thus deduce

Received September 26, 1950.

that $k + y$, and hence k, is algebraic over P. In short, K is algebraic over P. If P has prime characteristic, we have reached the other possibility stated in the conclusion of the lemma, so it remains only to exclude the possibility that P has characteristic 0 (which means that it is the field of rational numbers). This we do as follows. For any integer i we have an expression like (1) for $y + i$:

$$(2) \qquad y + i = (1 - \eta_i)(\eta_i - \epsilon_i)^{-1}.$$

Moreover, the definition of η_i and ϵ_i shows that they lie in the normal field, say Q, generated by y over P. But Q, being a finite-dimensional extension of P, contains only a finite number of roots of unity. This leaves us powerless to account for the infinite number of elements in (2).

Proof of the theorem. If $A \neq Z$, choose any element x not in Z, and let L be the field generated by Z and x. Then the hypothesis of the lemma is fulfilled (with Z playing the role of K). The possibility that Z has prime characteristic and is algebraic over its prime subfield is ruled out by the first theorem of Jacobson cited above. So it must be true that L is purely inseparable over Z. This is the case for every x, and we contradict the second theorem of Jacobson.

Theorem 7 of [1] actually states that a non-commutative division ring is generated by its nth powers. Our theorem can be given a corresponding extension as follows. For every x of a non-commutative division ring A, let there be given a positive integer $n(x)$ such that $n(x) = n(a^{-1}xa)$ for all $a \neq 0$; let B be the division subring generated by the elements $x^{n(x)}$; then $B = A$. For B is invariant under all inner automorphisms, and if $B \neq A$ then by the theorem of Cartan-Brauer-Hua [1, Theorem 2] B is contained in the centre of A, contradicting the above theorem.

In conclusion we discuss two possibilities of generalization. In the first place we might consider relaxing the requirement that A be a division ring. In fact, our theorem remains correct if we merely assume that A is semi-simple in the sense of Jacobson [2]. The manœuvre for proving this has become fairly standard since the appearance of Jacobson's paper. If P is a primitive ideal in A, our hypothesis is inherited by A/P; if we prove that each A/P is commutative we will know that A is commutative, and so we need only consider the case where A is primitive. We represent A as a dense ring of linear transformations in a vector space V over a division ring. We now in effect check our theorem for two-by-two matrices. In detail: if V is more than one-dimensional, let α and β be linearly independent vectors, and let x be an element of A sending α into itself and annihilating β. It is impossible for any power of x to be in the centre. So V is one-dimensional, and we are back to the division ring case of the theorem.

Another path along which to proceed is to have a polynomial more general than x^n. We shall not attempt more than the case where n is independent of

x, although it would be interesting to invent plausible "one-parameter families" generalizing $\{x^n\}$. We assume then that there exists a polynomial f with coefficients in Z (we can suppose it has no constant term) such that $f(x) \in Z$ for every x. Since A then satisfies the identity $f(x)y - yf(x) = 0$, it follows forthwith from [4, Theorem 1] that A is finite-dimensional over Z. But as a matter of fact it is again true that A is commutative. For suppose f has smallest possible degree among polynomials with $f(x) \in Z$. We can suppose there is an element u in Z no power of which is 1 (otherwise Z would be of prime characteristic and algebraic over its prime field, etc.). Consider the polynomial $g(x) = f(x) - u^n f(xu^{-1})$, n being the degree of f; the degree of g is less than n, and it again has the property $g(x) \in Z$ for every x. The only way out is for g to be identically zero, which means $f(x) = x^n$, and we are back to the old case.

One must step cautiously in attempting to generalize this last result beyond division rings: observe that the ring of two-by-two matrices over $GF(2)$ satisfies the identity $x^8 = x^2$.

References

[1] L. K. Hua, *Some properties of a sfield*, Proc. Nat. Acad. Sci. USA, vol. 35 (1949), 533-537.

[2] N. Jacobson, *The radical and semi-simplicity for arbitrary rings*, Amer. J. of Math., vol. 67 (1945), 300-320.

[3] ——— *Structure theory for algebraic algebras of bounded degree*, Ann. of Math., vol. 46 (1945), 695-707.

[4] I. Kaplansky, *Rings with a polynomial identity*, Bull. Amer. Math. Soc., vol. 54 (1948), 575-580.

University of Chicago

DISCRETE SEMI-ORDERED LINEAR SPACES

ISRAEL HALPERIN AND HIDEGORO NAKANO

1. Introduction. Let R be a *semi-ordered linear space*, that is, a *vector lattice* in Birkhoff's terminology [2]. An element $a \in R$ is said to be *discrete*, if for every element $x \in R$ such that $|x| \leqslant |a|$ there exists a real number α for which $x = \alpha a$. For every pair of discrete elements $a, b \in R$ we have $|a| \cap |b| = 0$ or there exists a real number α for which $b = \alpha a$ or $a = \alpha b$. Because, putting

$$c = |a| \cap |b|,$$

we have $c = \alpha a$, $c = \beta b$ for some real numbers α, β.

A system of elements $a_\lambda \in R (\lambda \in \Lambda)$ is said to be *complete*, if $|x| \cap |a_\lambda| = 0$ for all $\lambda \in \Lambda$ implies $x = 0$. R is said to be *universally continuous* if for every system of positive elements $a_\lambda \in R (\lambda \in \Lambda)$ there exists $\bigcap_{\lambda \in \Lambda} a_\lambda$ (*conditionally complete* in Birkhoff's terminology [2]).

Definition. A semi-ordered linear space R is said to be *discrete*, if R is universally continuous and has a complete system of discrete elements.

Let R be universally continuous. We shall use the notation $a_\lambda \downarrow_{\lambda \in \Lambda} a$ to mean: $a = \bigcap_{\lambda \in \Lambda} a_\lambda$ and for all $\lambda_1, \lambda_2 \in \Lambda$ there exists $\lambda \in \Lambda$ with $a_\lambda \leqslant a_{\lambda_1} \cap a_{\lambda_2}$. A linear functional L on R is said to be *universally continuous*, if

$$R \ni a_\lambda \downarrow_{\lambda \in \Lambda} 0 \quad \text{implies} \inf_{\lambda \in \Lambda} |L(a_\lambda)| = 0.$$

The totality of universally continuous linear functionals on R is said to be the *conjugate space* of R and denoted [5] by $\overline{R}$. R is said to be *semi-regular*, if R is universally continuous and $\bar{x}(a) = 0$ for all $\bar{x} \in \overline{R}$ implies $a = 0$.

Let R be semi-regular. A sequence of elements $a_\nu \in R$ $(\nu = 1, 2, \ldots)$ is said to be w-*convergent* to $a \in R$, if

$$\lim_{\nu \to \infty} \bar{x}(a_\nu) = \bar{x}(a) \qquad \text{for every } \bar{x} \in \overline{R}$$

and then we write $\text{w-}\lim_{\nu \to \infty} a_\nu = a$.

A sequence $a_\nu \in R$ $(\nu = 1, 2, \ldots)$ is said to be $|$w$|$-*convergent* to $a \in R$, if

$$\lim_{\nu \to \infty} \bar{x}(|a_\nu - a|) = 0 \qquad \text{for every } \bar{x} \in \overline{R},$$

and then we write $|\text{w}|\text{-}\lim_{\nu \to \infty} a_\nu = a$.

In a semi-ordered linear space R we have *order convergence*, i.e., we write $\lim_{\nu \to \infty} a_\nu = a$, if there exists a sequence of elements $R \ni l_\nu \downarrow_{\nu=1}^{\infty} 0$ such that

Received March 31, 1950.

$$|a_\nu - a| \leqslant l, \qquad (\nu = 1, 2, \ldots).$$

Kantorovitch [3] introduced *star convergence*, i.e., we write $\text{s-}\lim_{\nu\to\infty} a_\nu = a$ if every partial sequence from $a_\nu \in R$ ($\nu = 1, 2, \ldots$) contains a partial sequence which is order convergent to a.

We have furthermore *individual convergence* [7] i.e., we write $\text{ind-}\lim_{\nu\to\infty} a_\nu = a$, if

$$\lim_{\nu\to\infty} (a_\nu \frown x) \smile y = (a \frown x) \smile y \qquad \text{for all } x, y \in R;$$

and *star individual convergence*, i.e., we write $\text{s-ind-}\lim_{\nu\to\infty} a_\nu = a$ if every partial sequence from $a_\nu \in R(\nu = 1, 2, \ldots)$ contains a partial sequence which is individually convergent to a.

The purpose of this paper is to prove the

THEOREM. *Each of the following is necessary and sufficient in order that R should be discrete.*

(A) R is semi-regular and w-convergence coincides with |w|-convergence.
(B) R is semi-regular and star individual convergence coincides with individual convergence.
(C) R is semi-regular and |w|-convergence implies individual convergence.

The letters (A), (B), (C) will be used for reference throughout the paper, and R will denote a semi-ordered linear space.

2. LEMMA 1.[1] *If R is discrete, then R is semi-regular and* w*-convergence coincides with* |w|*-convergence, that is,*

$$\text{w-}\lim_{\nu\to\infty} x_\nu = 0 \quad \textit{implies} \quad \text{w-}\lim_{\nu\to\infty} |x_\nu| = 0.$$

Proof. If R is discrete, then R is universally continuous by definition. Furthermore R is semi-regular, because for every discrete element $a \neq 0$ we obtain a linear functional $\bar{a}$ in $\bar{R}$ as

$$[a]x = \bar{a}(x)a \qquad (x \in R)$$

for the projector (cf. [4]) $[a]$ of a.

Let $0 \leqslant a_\lambda \in R$ ($\lambda \in \Lambda$) be a complete system of discrete elements. Then we have obviously

$$\bigcap(1 - [a_{\lambda_1} + \ldots + a_{\lambda_\kappa}]) = 0$$

for all finite numbers of elements $\lambda_1, \ldots, \lambda_\kappa \in \Lambda$. Therefore we have by definition

$$\bigcap\bar{a}(1 - [a_{\lambda_1} + \ldots + a_{\lambda_\kappa}]) = 0$$

for every positive $\bar{a} \in \bar{R}$.

[1]From Lemma 1 we conclude immediately that in l_1 space weak convergence coincides with norm convergence, as was proved first by J. Schur [9].

We assume that $x_\nu \in R$ $(\nu = 1, 2, \ldots)$ is w-convergent to zero but not |w|-convergent to zero and derive a contradiction. We can suppose that for some positive $\bar{a} \in \overline{R}$ the inequality $\bar{a}(|x_\nu|) > 2$ holds for an infinite number of ν, hence

$$\bar{a}(x_\nu^+) > 1 \text{ or } \bar{a}(x_\nu^-) > 1$$

for an infinite number of ν. Replacing x_ν by $-x_\nu$ if necessary, we can suppose $\bar{a}(x_\nu^+) > 1$ for an infinite number of ν and hence (using only these x_ν) for all x_ν.

Now for each $\mu = 1, 2, \ldots$ define x_{ν_μ} and a projector

$$P_\mu = [a_{\mu_1} + \ldots + a_{\mu_\kappa}]$$

(with a finite number of indices $\mu_1, \ldots, \mu_\kappa \in \Lambda$, $\kappa = \kappa(\mu)$) by induction on μ so that:

$$\text{(i)} \qquad \bar{a}((\mathsf{U}_{\rho<\mu} P_\rho)|x_{\nu_\mu}|) < \tfrac{1}{3},$$

$$\text{(ii)} \qquad P_\mu \mathsf{U}_{\rho<\mu} P_\rho = 0,$$

$$\text{(iii)} \qquad \bar{a}((1 - \mathsf{U}_{\rho\leq\mu} P_\rho)|x_{\nu_\mu}|) < \tfrac{1}{3}.$$

Set $Q_\mu = [P_\mu x_{\nu_\mu}^+]$ and $Q = \mathsf{U}_{\mu=1}^\infty Q_\mu$. Then $\bar{a}Q$ is in $\overline{R}$, yet $\bar{a}Q(x_{\nu_\mu}) > \frac{1}{3}$ for all μ, contradicting the assumption $\lim_{\mu\to\infty} \bar{a}Q(x_{\nu_\mu}) = 0$.

LEMMA 2. *Let R be semi-regular. For a positive $p \in R$, if*

$$\text{w-}\lim_{\nu\to\infty} x_\nu = 0, \; |x_\nu| \leqslant p \qquad (\nu = 1, 2, \ldots)$$

implies $\text{w-}\lim_{\nu\to\infty} |x_\nu| = 0$, *then the normal manifold $[p]R$ is discrete.*

Proof. If $[p]R$ is not discrete, then there exists an element p_0 which we choose to denote also by $p(0, 1)$, such that $0 \neq [p_0] \leqslant [p]$, $[p_0]R$ has no discrete element except 0, and furthermore $[p_0]R$ is regular, i.e., there exists a positive $\bar{a} \in \overline{R}$ such that if $(0 \leqslant x \in R)$

$$\bar{a}(x) = 0 \quad \text{implies} \quad [p_0]\, x = 0.$$

For such a positive $\bar{a} \in \overline{R}$, we see easily that there exist two elements $p(0, 2^{-1})$, $p(2^{-1}, 1)$ such that

$$[p_0] = [p(0, 1)] = [p(0, 2^{-1})] + [p(2^{-1}, 1)],$$
$$\bar{a}([p(0, 2^{-1})]p) = \bar{a}([p(2^{-1}, 1)]p).$$

Thus we obtain by induction elements

$$p(\mu 2^{-\nu}, (\mu + 1)2^{-\nu}) \qquad (\mu = 0, 1, 2, \ldots, 2^\nu - 1; \nu = 1, 2, \ldots)$$

such that

$$[p(\mu 2^{-\nu}, (\mu+1)2^{-\nu})]$$
$$= [p(2\mu 2^{-\nu-1}, (2\mu+1)2^{-\nu-1})] + [p((2\mu+1)2^{-\nu-1}, 2(\mu+1)2^{-\nu-1})],$$
$$\bar{a}([p(2\mu 2^{-\nu-1}, (2\mu+1)2^{-\nu-1})]p)$$
$$= \bar{a}([p((2\mu+1)2^{-\nu-1}, 2(\mu+1)2^{-\nu-1})]p).$$

Putting $x_\nu = \sum_{\mu=0}^{2^\nu-1} (-1)^\mu [p(\mu 2^{-\nu}, (\mu+1)2^{-\nu})]p$, we have

$$|x_\nu| = [p_0]p \qquad (\nu = 1, 2, \ldots)$$

and hence naturally

$$\lim_{\nu\to\infty} \bar{a}(|x_\nu|) = \bar{a}([p_0]p) \neq 0. \tag{2.1}$$

On the other hand we can prove

$$\lim_{\nu\to\infty} \bar{b}(x_\nu) = 0 \qquad \text{for every } \bar{b} \in \bar{R}.$$

This can be done as follows: For a positive $\bar{b} \in \bar{R}$, define a function of a real variable $\bar{b}(t)$, $0 \leqslant t \leqslant 1$, by

$$\bar{b}(t) = \bar{b}(\bigcup_{\mu 2^{-\nu} \leqq t} [p((\mu-1)2^{-\nu}, \mu 2^{-\nu})]p).$$

Then it is not difficult to see that $\bar{b}(t)$ is absolutely continuous:

$$\bar{b}(t) = \int_0^t g(s)ds \qquad (0 \leqslant t \leqslant 1)$$

for some summable function $g(s)$. Now

$$\lim_{\nu\to\infty}\left(\sum_{\text{odd }\mu} \int_{\mu 2^{-\nu}}^{(\mu+1)2^{-\nu}} g(s)ds\right) = \lim_{\nu\to\infty}\left(\sum_{\text{even }\mu} \int_{\mu 2^{-\nu}}^{(\mu+1)2^{-\nu}} g(s)ds\right) = \frac{1}{2}\int_0^1 g(s)ds.$$

This is easily proved for continuous $g(s)$ and easily extended to all summable $g(s)$ (cf. [1]). Now the above shows that

$$\lim_{\nu\to\infty} \bar{b}(x_\nu^+) = \lim_{\nu\to\infty} \bar{b}(x_\nu^-)$$

and hence $\lim_{\nu\to\infty} \bar{b}(x_\nu) = 0$ for every positive $\bar{b} \in R$. Therefore we have w-$\lim_{\nu\to\infty} x_\nu = 0$ but not $|\mathrm{w}|$-$\lim_{\nu\to\infty} x_\nu = 0$ (by (2.1) contradicting the assumption.

In this proof, let y_γ $(\gamma = 1, 2, \ldots)$ be the sequence consisting of all elements

$$[p(\mu 2^{-\nu}, (\mu+1)2^{-\nu})]p \quad (\mu=0, 1, 2, \ldots, 2^\nu-1; \nu=1, 2, \ldots).$$

Then every partial sequence from $y_\gamma(\gamma = 1, 2, \ldots)$ contains a partial sequence $y_{\gamma_\nu}(\nu = 1, 2, \ldots)$ such that

$$\sum_{\nu=1}^{\infty} \bar{a}(y_{\gamma_\nu}) < +\infty.$$

Since $0 \leqslant y_{\gamma_\nu} \leqslant [p_0]p$ ($\nu = 1, 2, \ldots$), putting $y_0 = \limsup_{\nu\to\infty} y_{\gamma_\nu}$, we conclude that $\bar{a}(y_0) = 0$, and hence $y_0 = 0$. Therefore we have $\operatorname{s-lim}_{\gamma\to\infty} y_\gamma = 0$, while $\limsup_{\gamma\to\infty} y_\gamma = [p_0]p \neq 0$. Thus we obtain further:

LEMMA 3. *If R is semi-regular and for a positive $p \in R$ if*

$$\operatorname{s-lim}_{\nu\to\infty} x_\nu = 0, \ |x_\nu| \leqslant p \qquad (\nu = 1, 2, \ldots)$$

implies $\lim_{\nu\to\infty} x_\nu = 0$, *then the normal manifold $[p]R$ is discrete.*

Conversely we have

LEMMA 4. *If R is discrete, then*

$$\operatorname{s-ind-lim}_{\nu\to\infty} x_\nu = 0 \textit{ implies } \operatorname{ind-lim}_{\nu\to\infty} x_\nu = 0.$$

Proof. Let $a_\lambda (\lambda \in \Lambda)$ be a complete system of discrete elements. If

$$\operatorname{s-lim}_{\nu\to\infty} x_\nu = 0, \ |x_\nu| \leqslant p \qquad (\nu = 1, 2, \ldots),$$

then we have obviously

$$\lim_{\nu\to\infty} [a_\lambda]|x_\nu| = 0 \qquad \text{for every } \lambda \in \Lambda.$$

Putting $x_0 = \limsup_{\nu\to\infty} |x_\nu|$, we have

$$[a_\lambda]x_0 = \limsup_{\nu\to\infty} [a_\lambda]\,|x_\nu| = 0 \qquad \text{for every } \lambda \in \Lambda.$$

Since $a_\lambda (\lambda \subset \Lambda)$ is a complete system in R, we obtain then $x_0 = 0$. Therefore we have $\lim_{\nu\to\infty} x_\nu = 0$.

By virtue of Lemmas 1 and 2 we have: *the condition* (A) *is necessary and sufficient in order that* R *be discrete.* And furthermore, as an immediate consequence from Lemmas 3 and 4 we have: *the condition* (B) *is necessary and sufficient in order that* R *be discrete.*

Since $\operatorname{s-lim}_{\nu\to\infty} x_\nu = 0$ implies $|\mathrm{w}|\text{-}\lim_{\nu\to\infty} x_\nu = 0$, as can be seen from the definitions, we obtain by Lemma 3:

LEMMA 5. *Let R be semi-regular. For a positive $p \in R$ if*

$$|\mathrm{w}|\text{-}\lim_{\nu\to\infty} x_\nu = 0, \ |x_\nu| \leqslant p \qquad (\nu = 1, 2, \ldots)$$

implies $\lim_{\nu\to\infty} x_\nu = 0$, *then the normal manifold $[p]R$ is discrete.*

LEMMA 6. *If R is discrete, then*

$$|\mathrm{w}|\text{-}\lim_{\nu\to\infty} x_\nu = 0 \textit{ implies } \operatorname{ind-lim}_{\nu\to\infty} x_\nu = 0.$$

Proof. It is sufficient to prove this for the case $x_\nu \geqslant 0$ $(\nu = 1, 2, \ldots)$. Now for fixed $p \geqslant 0$, let

$$x_p{}^* = \limsup_{\nu\to\infty} (x_\nu \frown p).$$

We need only prove that $x_p{}^* = 0$ for each $p \geqslant 0$. But for every discrete element $a \in R$ and any $\bar{a} \in \bar{R}$, it is easy to prove that $\bar{a}([a]x_p{}^*) = 0$. Hence $[a]x_p{}^* = 0$ for every discrete $a \in R$, implying that $x_p{}^* = 0$ as required.

By virtue of Lemmas 5 and 6 we obtain: *the condition* (C) *is necessary and sufficient in order that R be discrete.*

Remark 1. We can also prove the theorem algebraically without the use of classical integration theory (see [6]), if we apply some results obtained in an earlier paper [8].

Remark 2. The theorem is also valid with the following definition: R is discrete, if R is continuous and has a complete system of discrete elements, replacing the condition that R is semi-regular by the conditions that R is continuous and to every element $p \neq 0$ there exists $q \neq 0$ such that $[q] \leqslant [p]$ and $[q]R$ is regular.

References

[1] C. R. Adams and A. P. Morse, *Random sampling in the evaluation of a Lebesgue integral*, Bull. Amer. Math. Soc., vol. 45 (1939), 442-447.

[2] G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Coll. Pub. 25 (1940).

[3] L. Kantorovitch, *Lineare halbgeordnete Räume*, Math. Sbornik, vol. 2 (44) (1937), 121-168.

[4] H. Nakano, *Teilweise geordnete Algebra*, Jap. J. Math., vol. 17 (1941), 425-511.

[5] ———, *Stetige lineare Funktionale auf dem teilweisegeordneten Modul*, J. Fac. Sci. Imp. Univ. Tokyo, vol. 4 (1942), 201-382.

[6] ———, *Discrete semi-ordered linear spaces* (in Japanese), Functional Analysis, vol. 1 (1947-9), 204-207.

[7] ———, *Ergodic theorems in semi-ordered linear spaces*, Ann. Math., vol. 49 (1948), 538-556.

[8] ———, *Modulared semi-ordered linear spaces*, Tokyo Math. Book Series, vol. 1 (Tokyo, 1950), §26, §27.

[9] J. Schur, *Über lineare Transformationen in der Theorie der unendlichen Reihen*, J. f. reine u. angew. Math., vol. 151 (1921), 79-111.

Queen's University and
Tokyo University

A GENERALIZATION OF THE PAPPUS CONFIGURATION

GERALD BERMAN

1. Introduction. A configuration is a system of m points and n lines such that each point lies on μ of the lines and each line contains ν of the points. It is usually denoted by the symbol (m_μ, n_ν), with $m\mu = n\nu$. Two configurations corresponding to the same symbol are said to be equivalent if there exist 1–1 mappings of the points and lines of one onto the points and lines of the other which preserve the incidence relations. It is a combinatorial problem to determine whether a given set of integers m, n, μ, ν with $m\mu = n\nu$ corresponds to an abstract configuration, and a geometric problem to determine whether the configuration exists in a given geometry. For example, there are two inequivalent configurations corresponding to the symbol $(12_4, 16_3)$, both of which exist in the real projective plane. A configuration is said to be inscriptible in a plane cubic if there exists an equivalent configuration whose points lie on the cubic. For such a configuration $\nu = 3$.

A family of configurations K_n corresponding to the symbol $(3n_n, n^2{}_3)$ $(n = 1, 2, \ldots)$ will be studied in this paper. K_1 is a line containing three distinct points, K_2 is the complete quadrilateral, K_3 is the Pappus configuration, and K_4 is a configuration studied by Zacharias [5]. In section **2** it will be shown that K_n contains configurations of the type K_q if n is a multiple of q. In section **3** it will be shown that K_n is inscriptible in the plane cubic curve as a real configuration with two degrees of freedom, and consequently exists in the real projective plane. This generalizes a result proved by Feld [2] for the Pappus configuration.

2. The family of configurations K_n. Let A_i, B_i, and C_i $(i = 0, 1, \ldots, n-1)$ be called points, and let (ij) $(i, j = 0, 1, \ldots, n-1)$ be called lines, where (ij) represents the triple of points A_i, B_j, C_k subject to the condition

$$i + j + k \equiv 0 \pmod{n}. \tag{2.1}$$

K_n is defined abstractly as the system of $3n$ points A_i, B_i, C_i $(i = 0, 1, \ldots, n-1)$ and n^2 lines (ij) $(i, j = 0, 1, \ldots, n-1)$. It can easily be verified that each of the $3n$ points lie on n of the lines, and each of the n^2 lines contains 3 of the points, so that the configuration has the symbol $(3n_n, n^2{}_3)$. The $3n$ points of K_n are the vertices of 3 n-gons in perspective in pairs from the vertices of the third, the n^2 lines of K_n being the lines of perspectivity. K_n can also be visualized as a $2n$-gon $A_0B_0A_1B_1 \ldots A_{n-1}B_{n-1}$ with the lines A_iB_j passing through the point C_k $(i + j \equiv -k \bmod n;\ k = 0, 1, \ldots, n-1)$.

If n is not a prime number, the configuration K_n has non-trivial components

Received April 25, 1950.

which are configurations belonging to the same family. In the proof of the following theorem the matrix (a_{ij}) $(i = 1, 2; j = 1, 2, \ldots, f)$ will represent the f^2 lines $(a_{1r}a_{2s})$ $(r, s = 1, 2, \ldots, f)$.

THEOREM 2. 1. *If n is a multiple of q, K_n contains $(n/q)^2$ distinct configurations K_q no two of which have a line in common. Each line of K_n is a line of one of the K_q, and each point of K_n is a point of n/q of the K_q.*

Consider the r^2 matrices

$$K(ij) \equiv \begin{pmatrix} i & r+i & 2r+i & \ldots & (q-1)r+i \\ j & r+j & 2r+j & \ldots & (q-1)r+j \end{pmatrix} \quad (i, j = 0, 1, \ldots, r-1)$$

where $r = n/q$. The lines represented by $K(ij)$ are the lines of a K_q for all $i, j = 0, 1, \ldots, r-1$. To see this define

2.2 $$A_{kr+i} \equiv A^*_k, \quad B_{kr+j} \equiv B^*_k, \quad C_{kr-i-j} \equiv C^*_k \qquad (k = 0, 1, \ldots, q-1).$$

The $3q$ points 2.2 are the only points on the lines represented by $K(ij)$. From the condition 2.1 for collinearity it follows that the points A^*_k, B^*_l, C^*_m will be collinear if and only if

2.3 $$r(k + l + m) \equiv 0 \qquad (\mathrm{mod}\ n).$$

Since $rq = n$, 2.3 holds if and only if

2.4 $$k + l + m \equiv 0 \qquad (\mathrm{mod}\ q).$$

The points 2.2 and the lines A^*_k, B^*_l, C^*_m subject to the condition 2.4 form a K_q by definition.

The r^2 configurations K_q represented by $K(ij)$ $(i, j = 0, 1, \ldots, n-1)$ are all distinct. By a consideration of the matrices $K(ij)$ it is seen that no two have a line in common. Furthermore, any point of K_n occurs in exactly r of the K_q. The $q^2r^2 = n^2$ lines of the r^2 K_q make up all the lines of K_n.

COROLLARY 1. *The 16 lines of K_4 can be divided into four sets of four lines which form complete quadrilaterals.*

This result was obtained by Zacharias [5].

COROLLARY 2. *K_{3q} contains q^2 distinct Pappus configurations.*

3. The inscription of K_n in the non-singular plane cubic curve. Any real non-singular cubic $\mathfrak{C}$ may be transformed into the Weierstrass canonical form by a suitable choice of the triangle of reference. Then the co-ordinates of any point on $\mathfrak{C}$ can be expressed parametrically in the form $(\wp u, \wp' u, 1)$ where $\wp u$ is the Weierstrass elliptic function. The point having the parameter u will be denoted by u. The necessary and sufficient condition that the points u, v, w be collinear is that

3.1 $$u + v + w \equiv 0 \qquad (\mathrm{mod}\ 2\omega, 2\omega')$$

where 2ω and $2\omega'$ are the periods of $\wp u$. The real plane cubics fall into two classes, unipartite and bipartite, depending upon whether they have one or two real circuits. For the bipartite cubic 2ω and $2\omega'/i$ are real and positive, while for the unipartite cubics 2ω and $2\omega'$ are conjugate complex. The points on the even branch of the bipartite cubic are given by values of the parameter of the form $u + \omega'$ where u is real. Points on the odd branch of either type are given by real values of the parameter.

The conditions that the $3n$ points

3.2 $$A_i, B_i, C_i \qquad (i = 0, 1, \ldots, n-1)$$

of $\mathfrak{C}$ should be points of a K_n are

3.3 $$A_i + B_j + C_k \equiv 0 \qquad (\text{mod } 2\omega, 2\omega')$$

with

3.4 $$i + j + k \equiv 0 \qquad (\text{mod } n).$$

Sum those equations of 3.3 having A_i in common:

$$\sum_{j,k=1}^{n-1} (A_i + B_j + C_k) \equiv 0 \qquad (\text{mod } 2\omega, 2\omega')$$

so that

$$n\,A_i \equiv -\sum_{j=0}^{n-1} (B_j + C_j) \qquad (\text{mod } 2\omega, 2\omega')$$

for $i = 0, 1, \ldots, n-1$. Thus

3.5 $$n\,A_0 \equiv n\,A_1 \equiv \ldots \equiv n\,A_{n-1} \qquad (\text{mod } 2\omega, 2\omega').$$

Similarly

3.6 $$nB_0 \equiv nB_1 \equiv \ldots \equiv nB_{n-1} \qquad (\text{mod } 2\omega, 2\omega'),$$

3.7 $$nC_0 \equiv nC_1 \equiv \ldots \equiv nC_{n-1} \qquad (\text{mod } 2\omega, 2\omega').$$

The equation $nu \equiv v$ (mod $2\omega, 2\omega'$) has n^2 distinct solutions

3.8 $$u \equiv v/n + 2(r\omega + s\omega')/n \quad (\text{mod } 2\omega, 2\omega') \quad (r, s = 0, 1, \ldots, n-1).$$

If $\mathfrak{C}$ is unipartite and v real, u will be real if and only if $r = s$. This leaves n distinct real solutions

$$u \equiv v/n + 2r(\omega + \omega')/n \quad (\text{mod } 2\omega, 2\omega_0) \qquad (r = 0, 1, \ldots, n-1).$$

Thus, since the points A_i $(i = 0, 1, \ldots, n-1)$ are all distinct and since 3.5 holds we may take

3.9 $$A_i \equiv A + 2i(\omega + \omega')/n \quad (\text{mod } 2\omega, 2\omega') \qquad (i = 0, 1, \ldots, n-1),$$

with A real. Similarly we may take

3.10 $\quad B_i \equiv B + 2i(\omega + \omega')/n \pmod{2\omega, 2\omega'} \qquad (i = 0, 1, \ldots, n-1),$

3.11 $\quad C_i \equiv C + 2i(\omega + \omega')/n \pmod{2\omega, 2\omega'} \qquad (i = 0, 1, \ldots, n-1),$

with B, C real. The condition 3.3 will be satisfied by the points 3.2 if and only if

3.12 $$A + B + C \equiv 0 \pmod{2\omega, 2\omega'}.$$

The configuration will degenerate if any two of the sets of points 3.9, 3.10. 3.11 are the same. Thus nA, nB and nC must be different modulo 2ω, $2\omega'$.

If $\mathfrak{C}$ is bipartite and v real, u will be real if and only if $s = 0$. This leaves n distinct real solutions

3.13 $\quad u \equiv v/n + 2r\omega/n \pmod{2\omega, 2\omega'} \qquad (r = 0, 1, \ldots, n-1).$

Thus we may take

3.14 $\quad A_i \equiv A + 2i\omega/n \pmod{2\omega, 2\omega'} \qquad (i = 0, 1, \ldots, n-1)$

3.15 $\quad B_i \equiv B + 2i\omega/n \pmod{2\omega, 2\omega'} \qquad (i = 0, 1, \ldots, n-1)$

3.16 $\quad C_i \equiv C + 2i\omega/n \pmod{2\omega, 2\omega'} \qquad (i = 0, 1, \ldots, n-1)$

with A, B, and C real satisfying condition 3.12. Thus nA, nB, and nC must be different, as before, so that the configuration will not degenerate.

If $\mathfrak{C}$ is bipartite and $u - \omega'$ real, then the points 3.13 will all be real and on the even branch. Thus if any one of the points 3.14 lies on the even branch, i.e. if $A - \omega'$ is real, all the points 3.14 lie on the even branch. A similar statement holds for the points 3.15 and 3.16. By condition 3.12 which must be satisfied by the points of K_n, either none or exactly two of the sets of points 3.14, 3.15, 3.16 lie on the even branch.

We have proved

THEOREM 3.1. *K_n may be inscribed in a non-singular plane cubic $\mathfrak{C}$ with two degrees of freedom. Any two real points u, v such that nu, nv, and $-n(u + v)$ are different* (mod 2ω, $2\omega'$) *may be selected as a pair of points of the configuration, and the remaining points are uniquely determined. If $\mathfrak{C}$ is bipartite the $3n$ points of K_n fall into three sets of n points such that either two or none of the sets lie on the even branch.*

We have also proved

THEOREM 3.2. *K_n exists in the real projective plane for all n.*

REFERENCES

[1] H. S. M. Coxeter, *Self-dual configurations and regular graphs*, Bull. Amer. Math. Soc., vol. 56 (1950), 413-455.

[2] J. M. Feld, *Configurations inscriptible in a plane cubic curve*, Amer. Math. Monthly, vol. 43 (1936), 549-555.

[3] O. Hesse, *Über Curven dritter Ordnung und die Kegelschnitte welche diese Curven in drei verschiedenen Punkten berühren*, J. Reine Angew. Math., vol. 36 (1848), 143-176.
[4] Th. Reye, *Das Problem der Configurationen*, Acta Math., vol. 1 (1882), 97-108.
[5] M. Zacharias, *Untersuchungen über ebene Konfigurationen* (12_4, 16_3), Deutsche Math., vol. 6 (1941), 147-170.
[6] ——, *Eine ebene Konfiguration* (12_4, 16_3) *in der Dreiecksgeometrie*, Monatsh. Math. Phys., vol. 44 (1936), 153-158.
[7] ——, *Über den Zusammenhang des Morleyschen Satzes von den winkeldrittelnden Eckenlinienen eines Dreiecks mit den trilinearen Verwandtschaften in Dreieck und mit einer Konfiguration* (12_4, 16_3) *der Dreiecksgeometrie*, Deutsche Math., vol. 3 (1938), 36-45.

Illinois Institute of Technology

A NOTE ON SOME PERFECT SQUARED SQUARES

T. H. WILLCOCKS

1. Introduction. In a recent paper [5], general methods were described for the dissection of a square into a finite number n of unequal non-overlapping squares. In this note, examples of such perfect squares are given in which the sides and elements are relatively small integers; in particular, a dissection of a square into 24 different elements, which is believed to be the squaring of least order known at the present time. All the dissections which follow make use of auxiliary rectangles; that is to say, the squarings are compound. The following terminology, introduced by the authors of [2], will be used. A *dissection* of a rectangle R into a finite number n of non-overlapping squares is called a *squaring* of R of *order* n; and the n squares are the *elements* of the dissection. If the elements are all unequal and $n > 1$, the squaring is *perfect* and R is a *perfect rectangle*. A squared rectangle which contains a smaller squared rectangle is called *compound*, all others being *simple*. Two squared rectangles which have the same shape (i.e. proportional sides) but are not merely rigid displacements of each other are called *conformal;* two conformal rectangles are said to be *totally different* if C_2 times an element of the first is never equal to C_1 times an element of the second, where C_1 and C_2 are their respective corresponding sides.

Complete dissections will be expressed in the notation of C. J. Bouwkamp [3], which consists in enclosing within brackets the lengths of the sides of those squares whose upper sides lie in the same horizontal segment, the brackets being read in order from the top to the bottom of the rectangle.

2. Methods for constructing perfect squares. The following methods are based on the combination in various ways of certain related pairs of rectangles. A squared rectangle of sides x and y will be written (x, y) and if the squaring is perfect will be denoted by P (x, y). A rectangle which has but two elements equal, one (or both) of which is in a corner, will be called *quasi-perfect* and written Q (x, y). The imperfection of such rectangles may be either trivial or non-trivial, two equal squares constituting a *trivial* imperfection if they extend between the same two horizontal (or vertical) segments in the dissection [2, (5.22)]. In what follows, when reference is made to two perfect rectangles, it is to be understood that they are totally different, unless the contrary is stated.

2.1. *From two rectangles* P (x, y) *and two squares* x *and* y. An example of order 28 and reduced side 1015 is given in [1]. One of order 34 and reduced side 960 is derived from two P (479, 481). It is (218, 124, 137, 481), (49, 53, 22),

Received August 27, 1950.

(9, 128), (31), (45, 4), (88), (263), (216), (479, 240, 241), (239, 1), (76, 67, 99), (11, 24, 32), (59, 15, 2), (13), (44, 8), (139), (103).

2.11. *From two rectangles* P (x, y) *and two squares* $x - A$ *and* $y - A$ *where* A *is a corner element of one* P (x, y). In the resulting squaring the element A is overlapped. In favourable cases eight different perfect squares may be formed, but some arrangements of the border squares of the auxiliary rectangles yield a lesser number. Examples of order 33 and reduced sides 821, 823, 857, 861 and 884 are derived from two P (479, 481). These rectangles are (218, 124, 137), (49, 53, 22), (9, 128), (31), (45, 4), (88), (263), (216) and (76, 67, 99, 239), (11, 24, 32), (59, 15, 2), (13), (44, 8), (139), (103), (241, 1), (240) and the P-squares are obtained by making A successively, 139, 137, 103, 99, 76. (As the second of these rectangles is compound, reorientation of its auxiliary rectangle is necessary to make the squares 76, 99, 103, 139 successively occupy a corner.)

Note (a). This method fails if both rectangles are such that no two adjacent sides can be chosen each having more than two elements, e.g., the rectangles XIII, 1015 *g* and *h* in the classification of [3].

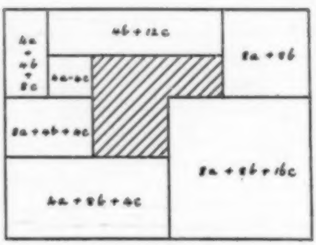

FIG. A

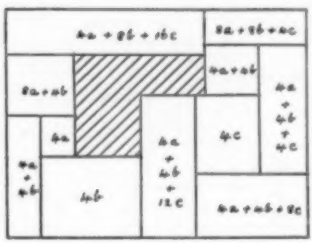

FIG. B

Note (b). This method can be used however if (i) the two rectangles P(x, y) have one corner element A in common. The structure of a square of 39th order of this type was given in [2], or if (ii) a P-rectangle and a Q-rectangle are used instead of two P-rectangles. An example of order 24 and reduced side 175 is (55, 39, 81), (16, 9, 14), (4, 5), (3, 1), (20), (56, 18), (38), (30, 51), (64, 31, 29), (8, 43), (2, 35), (33). One of order 26 and reduced side 492 is (57, 59, 56, 95, 225), (3, 14, 39), (55, 2,), (53, 11), (25), (17, 142), (125), (24, 60, 141), (255, 12), (36), (96), (15, 126), (111). The first of these was published in [4].

2.2. *From rectangles* P(x, y) *and* P$(x, x+y)$ *and a square* y. An example of order 26 and reduced side 608 was given in [2].

2.21. *From rectangles* P(x, y) *and* P$(x, x+y)$ *and squares* $x, x-A, x+y-A$ *where* A *is a corner element of a rectangle, and is overlapped as before.* Using the auxiliary rectangles employed in the last mentioned square of side 608, and

making A successively 136, 118, 113 and 95, squares of order 27 and reduced sides 849, 867, 872 and 890 are obtained.

2.3. *From rectangles* (x, y) *and* $(x+y, x+2y)$ *and squares* $y, x+y, x+y-A$ *and* $x+2y-A$. An example of order 28 and reduced side 577 is derived from Q (113, 127) and P(240, 353) and was first described in [4]. It is (224, 123, 129, 101); (28, 73), (117, 6), (111, 52,), (7, 66), (59), (113, 51, 21, 23, 16), (32, 337), (19, 2), (25), (11, 8), (65), (62), (240).

2.4. In the foregoing methods it has been necessary that the two auxiliary rectangles should have either no element in common, or one element of the one equal to a corner element of the other. In some cases, however, auxiliary rectangles not conforming to these conditions may serve, since certain P-rectangles (p, q) containing n elements can be transformed into Q-rectangles (p, q) of $n+4$ elements, sometimes in more than one way. It may happen, therefore, that two P-rectangles (x, y) may have one or more elements in common, but that one of them and a derived Q-rectangle may have none in common and method 2.11 may be applied to form a perfect square.

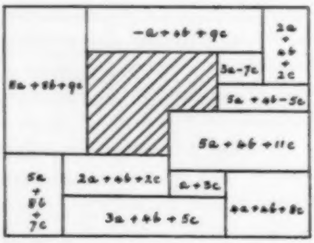

Fig. C

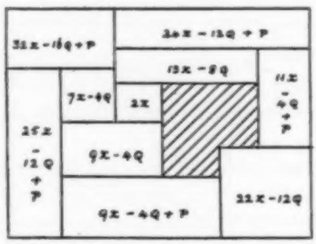

Fig. D

In the diagrams the shaded area represents a six-sided squared polygon, five angles of which are right angles. It will be seen that the three rectangles of figures A, B and C are conformal, whatever the structure of the polygon. In general A is a P-rectangle, B and C are Q-rectangles. There is a corner element $(4a + 4b + 8c)$ of A which is also a corner element of B and C, which fact allows further conformal pairs to be derived; for, if this element be removed from A and to the resulting polygon is added any arrangement of additional elements, then the same operation performed on B or C will result in a conformal figure.

Now suppose the corner square $(4a + 4b + 8c)$ of A be removed and four squares added so as to form the rectangle of Figure D, then a conformal rectangle may be formed in the following manner: the corner square $(8a + 8b + 16c)$ of A may be thought of as being composed of four equal smaller squares $(4a + 4b + 8c)$ containing a square x (of side zero). If the

new smaller corner square be removed and four squares added as before in corresponding positions, the square x will no longer (in general) vanish, and the conformal rectangle of Figure E will be formed.

An example follows of how such considerations may lead to a perfect square. The two P-rectangles (164, 118, 197), (39, 79,), (7, 32), (122, 49), (24, 8), (284), (73), (195), and (218, 124, 137), (49, 53, 22), (9, 128), (31), (45, 4), (88), (263), (216), are conformal and have one element (49) in common. The second of these rectangles is of the form D, and it is found that the conformal Q-rectangle of form E (274, 305, 465, 872), (243, 31), (212, 124), (88, 36), (274, 227), (543), (47, 180), (321), (1052), (864) together with the first of the P-rectangles (magnified) and the addition of squares 1642 and 1650 makes a perfect square of order 33 and reduced side 3566. An alternative dissection of this square of order 37 is (960, 964, 1642), (956, 4), (304, 268, 396), (44, 96, 128), (236, 60, 8,) (52), (176, 32), (556), (412), (305, 465, 872), (1650, 243, 31), (212, 124), (88, 36), (274, 227), (543), (47, 180), (321), (1052), (864).

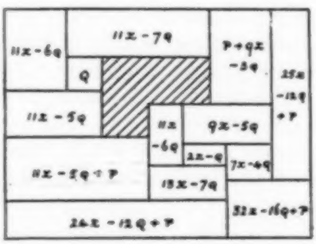

FIG. E

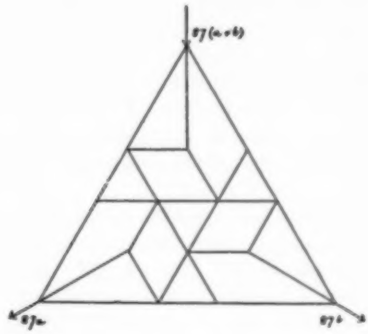

FIG. F

Now suppose two completely different squared polygons of the same size and shape are constructed by the method described in [2], and their shape is controlled so that with the addition of seven squares to one of them a P-rectangle such as Figure A above is formed, then Q-rectangles of the types B and C can be formed from the other. Since rectangle A can be combined with either B or C to form a square, the diagrams illustrated (or others derived from them) may be used to form non-trivially different squares of the same order and size (and these will, in general, be perfect.) Two such that have been evaluated are of order 73 and reduced side 1535484 and are constructed, by the methods of [2] from the diagram of Figure F, when a and b are assigned the values 506 and 185 respectively. This network was discussed in [2] and [3].

3. The construction of rectangles. The auxiliary rectangles employed in the squares described were discovered by constructing lists of isoperimetric rectangles of different sizes and selecting any pairs which were found to fulfil the required conditions. In the attempt to find P-squares of small reduced side, attention was centred mainly on semi-perimeters containing a number of small factors. It was shown in [2] that two rectangles which have the same c-net are isoperimetric.

References

[1] A. H. Stone, *Amer. Math. Monthly*, vol. 47 (1940), 570-572.

[2] R. L. Brooks, C. A. B. Smith, A. H. Stone, W. T. Tutte, *The dissection of rectangles into squares*, Duke Math. J., vol. 7 (1940), 312-340.

[3] C. J. Bouwkamp, *On the dissection of rectangles into squares* (Papers I, II, and III), Neder. Akad. Wetensch. Proc., vol. 49 (1946), 1176-1188 and vol. 50 (1947), 58-78.

[4] T. H. Willcocks, *Fairy Chess Review*, vol. 7, Aug./Oct. 1948.

[5] W. T. Tutte, *Squaring the square*, Can. J. Math., vol. 2 (1950), 197-209.

Bristol, England

MODULAR REPRESENTATIONS OF THE SYMMETRIC GROUP

J. H. CHUNG

Introduction. The theory of modular representations of the symmetric group was studied first by Nakayama (**5**, **6**), and later by Thrall and Nesbitt (**11**) and Robinson (**7**, **8**, **9**). Nakayama built up his elaborate theory of hooks for the express purpose of studying this problem, while Robinson's extensive work on the various phases of the relationship between Young diagrams, skew diagrams and star diagrams on the one hand, and representations of the symmetric group on the other, culminating in a set of relations among the degrees of the representations, serves as a starting point for this paper.

Brauer and Nesbitt (**2**) have shown in the general theory that, for a given prime p, the irreducible representations of a group may be separated into a number of p-blocks, each of which is characterized by the maximal power t of p which divides the degree of every representation of the block. If $g = p^a g'$, where g' is prime to p, then the equation $t + d = a$ relates t to the defect d of the block. If $t = a$, the block is of defect 0, while if $t = 0$, the block is of defect a. For the symmetric group Nakayama conjectured that the Young diagrams of all the representations of a single p-block had the same p-core after the removal of all their p-hooks. This conjecture was proved jointly in 1947 by Brauer and Robinson (**3**).

Brauer (**1**) also showed that the representations in a p-block of defect 1 can be arranged in a chain such that only adjacent members have a single modular component (with multiplicity 1) in common. For $n < 2p$, Nakayama succeeded in showing that in the case of the symmetric group S_n the ordering in the chain is precisely the natural order of the leg lengths r of the p-hooks, from $r = 0$ to $r = p - 1$, where each of the p distinct p-hooks is found in exactly one Young diagram.

The present paper extends Nakayama's result for blocks of defect 1 to values of $n \geqslant 2p$, and explains the derivation of a set of identities among the modular characters of the irreducible representations of a p-block of S_n. The nature of their linear dependence is studied in some detail. Notice is taken of the orthogonal relation between the coefficients in these identities and the columns of the matrix of decomposition numbers which gives the modular splitting of the irreducible representations of S_n, and this leads to an investigation of the nature of indecomposability in the regular representation of S_n. As a first step forward from Nakayama's one hook case, the indecomposables of the p-block of S_{2p} with zero p-core are obtained in a conclusive manner.

Received January 8, 1951. This paper contains the substance of a thesis prepared under the supervision of Professor G. de B. Robinson and submitted for a Ph.D. degree at the University of Toronto in May, 1950.

1. Background. Let p be a rational prime and $\mathfrak{p}$ be a fixed prime ideal divisor of p in an algebraic number field K. Suppose a group G is represented by matrices whose coefficients are taken from the ring $\mathfrak{o}$ of $\mathfrak{p}$-integers of K (i.e., numbers of the form α/β, where α and β are integers of K and β is prime to $\mathfrak{p}$), and let $Z_1, Z_2, \ldots, Z_k$ be the distinct irreducible representations of G. If we let $\bar{K}$ be the residue class field of $\mathfrak{o}$ (mod $\mathfrak{p}$) and replace every coefficient of the Z_i by its residue class (mod $\mathfrak{p}$), then the resulting modular representations $\bar{Z}_i$ will, in general, be reducible and will split into irreducible modular representations F_κ with coefficients in $\bar{K}$. The splitting may be denoted by

$$\bar{Z}_i = \sum_{\kappa=1}^{k^*} d_{i\kappa} F_\kappa \qquad (i = 1, 2, \ldots, k),$$

where $d_{i\kappa}$ is the multiplicity with which F_κ appears in $\bar{Z}_i$. These rational integers $d_{i\kappa} \geqslant 0$ are called the *decomposition numbers* (mod p) of G.

If $U_1, U_2, \ldots, U_{k^*}$ are the distinct indecomposable components of $\bar{R}$, the regular representation of G with entries in $\bar{K}$, then

$$U_\kappa = \sum_{\lambda=1}^{k^*} c_{\kappa\lambda} F_\lambda \qquad (\kappa = 1, 2, \ldots, k^*),$$

where the $c_{\kappa\lambda}$, rational integers $\geqslant 0$, are the Cartan invariants of G (for p), and are related to the decomposition numbers via the equations

$$c_{\kappa\lambda} = \sum_{i=1}^{k} d_{i\kappa} d_{i\lambda}.$$

There exists a representation (U_κ) of G in K which, if taken (mod p), becomes similar to U_κ. We then have

$$(U_\kappa) = \sum_{j=1}^{k} d_{j\kappa} Z_j,$$

and it is to these representations (U_κ) that we shall be referring (without ambiguity) as the *indecomposable representations* of G. Such an indecomposable representation has the property that its character vanishes for all elements of G whose orders are divisible by p, i.e., for all p-singular elements. In the case of the symmetric group S_n a p-regular element is simply a permutation the lengths of whose cycles are all prime to p, while a p-singular element has at least one cycle of length p or a multiple of p.

Corresponding to the foregoing relations among the representations we have character relations which are valid for the p-regular elements of G. If we denote by $\eta^{(\kappa)}$ the character of U_κ, by $\Phi^{(\kappa)}$ that of F_κ, and by $\xi^{(i)}$ that of Z_i, these relations are

1.1 $$\xi^{(i)} = \sum_{\lambda=1}^{k^*} d_{i\lambda} \Phi^{(\lambda)}$$

1.2 $$\eta^{(\kappa)} = \sum_{i=1}^{k} d_{i\kappa} \xi^{(i)}.$$

Regarding $(d_{i\kappa})$ as a matrix with i as row index and κ as column index, relations 1.1 and 1.2 indicate that the rows of this matrix give the splitting of the ordinary irreducible representations into their modular irreducible components, while the columns give the indecomposable representations of G in K as linear combinations of the ordinary irreducible representations with non-negative coefficients. It is this latter interpretation which will prove useful in the determination of the modular splitting of the irreducible representations of S_n.

A result in the modular theory which will also prove to be particularly useful is embodied in the following two formulae of Nakayama (**2**, p. 582):

1.3 $$\bar{\eta}^{(\kappa)*} = \sum_{\lambda} a_{\kappa\lambda}\, \eta^{(\lambda)} \qquad \text{(for } p\text{-regular elements of } G),$$

1.4 $$\varphi^{(\lambda)} = \sum_{\kappa} a_{\kappa\lambda}\, \bar{\varphi}^{(\kappa)} \qquad \text{(for } p\text{-regular elements of } H).$$

Observe that the same coefficients $a_{\kappa\lambda}$, which are positive integers or zero, appear in both formulae. The first states, in the notation of characters, that the representation of G induced by an indecomposable representation of a subgroup H of G, can be expressed as a linear combination of indecomposable representations of G, while the second states that if we restrict our attention to the element of a subgroup H of G, any modular irreducible representation of G becomes equivalent to a sum of modular irreducible representations of H.

A. Young showed that there exists a one-to-one correspondence between the irreducible representations of the symmetric group S_n and his tableaux or "diagrams", so that the same symbol can be used interchangeably for a Young diagram and for the corresponding irreducible representation. A generalization of the notion of a Young or *right* diagram is a *skew* diagram $[\alpha] - [\beta]$, introduced by Robinson (**9**), which consists of the nodes left after removing from the corner of a Young diagram $[\alpha]$ nodes which themselves make up a Young diagram $[\beta]$. If the skew diagram consists of disjoint constituents with no row or column in common, it is called a *disjoint* diagram. To every such diagram corresponds an induced representation; of particular significance are disjoint diagrams whose constituents are right diagrams. If, for example, there are two constituent right diagrams $[\beta]$ and $[\gamma]$, where $[\beta]$ is a representation of S_l and $[\gamma]$ a representation of S_m, then the resulting Kronecker product representation of the subgroup $S_l \times S_m$ is written $[\beta] \times [\gamma]$, and the representation of the symmetric group S_{l+m} on $l + m$ distinct symbols induced by this Kronecker product representation is written $[\beta] \cdot [\gamma]$. It is to $[\beta] \cdot [\gamma]$ that the forementioned diagram corresponds, and its reduction into irreducible components $[\alpha]$ of S_n $(n = l + m)$ takes the form

$$[\beta] \cdot [\gamma] = \sum_{\alpha} {}_{\alpha}\lambda_{\beta}{}^{\gamma} [\alpha].$$

The ${}_{\alpha}\lambda_{\beta}{}^{\gamma}$ are obtained via the Littlewood-Richardson rule (**4**, p. 119) for writing down the irreducible components of $[\beta] \cdot [\gamma]$.

It seems unnecessary to summarize here the theory of *hooks* as developed

in the papers of Nakayama and Robinson already referred to. We note in conclusion a recent paper by Nakayama and Osima (Nagoya Math. J. vol. 2 (1951), 111-117) in which an alternative proof is given of Nakayama's conjecture (cf. **3**).

2. The removal of the restriction $n<2p$ in Nakayama's one hook case. In studying the modular splitting of ordinary irreducible representations of S_n via his notion of a hook, Nakayama naturally started with the case of p-cores of n nodes and immediately reached the conclusion that the corresponding irreducible representations were also modular irreducible and that each of them formed by itself a block of defect 0. Further study along these lines led him to the result that each block of defect 1 contained exactly p representations, namely, those having Young diagrams with a given p-core of $n-p$ nodes and p-hooks of leg lengths $0, 1, 2, \ldots, p-1$, respectively—a result that he was able to prove only for the case $n < 2p$, but which followed directly for all n as soon as his conjecture (**6**, p. 423) was proved. For such a block he stated the following theorem (re-phrased):

2.1. *Let T_0 be a p-core of $n-p$ nodes and let $T_{0,r}$ be the* (*unique*) *diagram of n nodes with p-core T_0 and one p-hook of leg length r. Then the irreducible representation $[\beta]_r$ of S_n associated with $T_{0,r}$ possesses exactly one irreducible modular component* (*with multiplicity* 1) *in common with $T_{0,r+1} (r \neq p-1)$, one in common with $T_{0,r-1}$ $(r\neq 0)$, and none in common with $T_{0,s}$ $(s \neq r-1, r+1)$.*

To prove this theorem for $n<2p$, Nakayama utilized a result of Brauer (**1**) in the general modular theory concerning the arrangement in a chain of the representations in a block of defect 1, in which only neighbouring representations have a modular component (with multiplicity 1) in common, in order to identify his diagrams with the corresponding representations in the chain. His reason for considering values of $n < 2p$ was simply that only in this range could he be certain of having to contend only with blocks of defects 0 and 1.

To prove the theorem for all values of n, we accept the truth of the result for $n = p$ (Nakayama's proof covers this value), i.e., for a p-core of zero nodes. This means that the portion D_p of the D-matrix appropriate to the corresponding p-block of S_p is of the form

$$\begin{array}{c} [\alpha]_0 \\ [\alpha]_1 \\ [\alpha]_2 \\ \cdot \\ [\alpha]_{p-2} \\ [\alpha]_{p-1} \end{array} \begin{bmatrix} 1 & 0 & 0 \ldots 0 & 0 \\ 1 & 1 & 0 \ldots 0 & 0 \\ 0 & 1 & 1 \ldots 0 & 0 \\ \cdot & \cdot & \cdot \quad\; \cdot & \cdot \\ 0 & 0 & 0 \ldots 1 & 1 \\ 0 & 0 & 0 \ldots 0 & 1 \end{bmatrix} : D_p ,$$

where $[\alpha]_r$ is the representation of the block whose Young diagram is a p-hook of leg length r. The columns of D_p give the indecomposable components of

the regular representation that belong to the block, so that in particular $[\alpha]_r$ appears in two indecomposable components, namely $[\alpha]_{r-1} + [\alpha]_r$ and $[\alpha]_r + [\alpha]_{r+1}$.

Let $[\beta_0]$ be the representation of S_{n-p} $(n \geqslant p + 1)$ whose Young diagram is the p-core T_0, so that $[\beta_0]$ is modular irreducible and also forms by itself a p-block of S_{n-p} of defect 0. Denote by $[\beta]_r$ and $[\beta]_{r+1}$ the two representations of the p-block of S_n of defect 1 with p-core $[\beta_0]$ whose Young diagrams have p-hooks of leg lengths r and $r + 1$ respectively. In order to show that $[\beta]_r + [\beta]_{r+1}$ is an indecomposable component of the regular representation of S_n, we prove first a preliminary lemma.

2.2. *If $[\beta_0]$ is a p-core of $n - p$ nodes, and $H_r = [p - r, 1^r]$ a p-hook of leg length r, then of all the diagrams $[\beta]$ of n nodes that can be obtained by building H_r on $[\beta_0]$ in accordance with the Littlewood-Richardson rule, there is only one which has $[\beta_0]$ as p-core, namely the (unique) diagram which contains the p-hook H_r and p-core $[\beta_0]$.*

Proof. Nakayama demonstrated the existence of exactly one diagram $[\beta]_r$ of n nodes which possesses the desired p-hook and p-core, but we need to show that it actually arises as a result of building in accordance with the Littlewood-Richardson rule. The nodes of the p-hook in question can be thought of as being added along the rim of the p-core $[\beta_0]$ so as to form a skew hook equivalent to the right hook H_r, and the only point that needs verification is that this building on $[\beta_0]$ does not violate any of the restrictions laid down in the Littlewood-Richardson rule.

We observe that the first and last nodes of a skew hook (starting from the top right and going to the bottom left) correspond respectively to the head and foot of the equivalent right hook H_r, so that there are exactly as many rows $(r + 1)$ and columns $(p - r)$ represented in the skew hook as in the right hook. Then, since no two added symbols from a given row of H_r may appear in the same column of the resultant diagram $[\beta]$, the $p - r$ nodes in the first row of H_r must be assigned, in order, one to each column of the skew hook. Likewise, since each node of the first column of H_r must appear in a later row of $[\beta]$ than its predecessors of that column, the nodes of the first column of H_r must be assigned, in order, one to each row of the skew hook.

Suppose that we designate the nodes of the right hook H_r in the following way:

$$\begin{array}{llll} C_1 & C_2 & C_3 \ldots & C_s \\ R_2 & & & \\ R_3 & & & \\ \cdot & & & \\ \cdot & & & \\ R_{r+1} & & & \end{array}$$

where $s = p - r$ and $C_1 \equiv R_1$. Now consider the skew hook obtained by building on $[\beta_0]$ with H_r; from the preceding paragraph it is clear that there will be exactly one R per row ($C_1 \equiv R_1$) and exactly one C per column. Since none of the restrictions involved in the Littlewood-Richardson rule have been violated, $[\beta]_r$ can thus be obtained from this building process. To show that it appears only once, we notice that in the skew hook neither the C's nor the R's can be interchanged among themselves without violating the rule. Further no C can be interchanged with an R, for otherwise we would have two C's in a column of the product diagram. Finally $[\beta]_r$ is the only diagram which contains H_r as a p-hook and $[\beta_0]$ as its p-core.

To show that the process does not yield a diagram $[\beta]_t$ containing a p-hook H_t $(t \neq r)$ and p-core $[\beta_0]$, we observe that such a diagram $[\beta]_t$ will contain a skew hook $\overline{H}_t$ equivalent to H_t, containing $t + 1$ rows and $p - t$ columns. If such a skew hook is to arise from building on $[\beta_0]$ with H_r, we shall have more than one C in at least one column of $\overline{H}_t$ if $t > r$, and more than one R in at least one row of $\overline{H}_t$ if $t < r$. In either case a restriction in our building process is violated, and hence such a diagram $[\beta]_t$ cannot arise. This proves the lemma.

Proof of 2.1 *for* $n > p$: Since $[\alpha]_r + [\alpha]_{r+1}$ is an indecomposable representation of S_p (mod p), and $[\beta_0]$ (a p-core) is a modular irreducible representation of S_{n-p}, then

$$([\alpha]_r + [\alpha]_{r+1}) \times [\beta_0]$$

is an indecomposable Kronecker product representation of the direct product subgroup $S_p \times S_{n-p}$ of S_n. Further, by Nakayama's formula (1.3), the corresponding induced representation of S_n

$$[\alpha]_r \cdot [\beta_0] + [\alpha]_{r+1} \cdot [\beta_0]$$

is a sum of indecomposable representations of S_n, whose irreducible components are obtained via the Littlewood-Richardson rule applied to the induced representations $[\alpha]_r \cdot [\beta_0]$ and $[\alpha]_{r+1} \cdot [\beta_0]$. Now the only components that we are interested in are those that belong to the p-block with p-core $[\beta_0]$, and the preliminary lemma tells us that there will be exactly two such irreducible representations, one obtained from $[\alpha]_r \cdot [\beta_0]$ and the other from $[\alpha]_{r+1} \cdot [\beta_0]$. Denoting the representations or the corresponding Young diagrams by $[\beta]_r$ and $[\beta]_{r+1}$, we observe that $[\beta]_r$ and $[\beta]_{r+1}$ each contain one p-hook, of leg length r and $r + 1$ respectively; hence, by the Murnaghan-Nakayama recursion formula, their characters cannot vanish for all elements of S_n of the type $P \,.\, V$, where P is a cycle of length p. Since the vanishing of the character

for all p-singular elements of S_n is a necessary condition for indecomposability, it follows that neither $[\beta]_r$ nor $[\beta]_{r+1}$ is an indecomposable representation. However, since the p-hooks in $[\beta]_r$ and $[\beta]_{r+1}$ have parities of opposite signs, the character of the sum $[\beta]_r + [\beta]_{r+1}$ vanishes for all p-singular elements; inasmuch as these are the only representations in the block under consideration, the sum $[\beta]_r + [\beta]_{r+1}$ must be an indecomposable representation of S_n $(r \neq p - 1)$. It follows in exactly the same way that $[\beta]_{r-1} + [\beta]_r$ is an indecomposable representation of S_n $(r \neq 0)$. Hence for any p-core $[\beta_0]$ the ordering of the representations in the associated p-block of S_n of defect 1, such that only adjacent representations have a modular component in common, is the same as in the case of the $[\alpha]$'s; i.e., the part of the D-matrix corresponding to this block is again D_p. This completes the proof.

Before proceeding to investigate the modular splitting of representations whose Young diagrams contain two or more p-hooks, we deduce in the next section a number of relations among the characters of any particular block, which hold for all p-regular elements. It is these relations which play a vital role in our subsequent analysis.

3. Character relations for p-regular elements of S_n. In (**9**) Robinson obtained some relations among the degrees of irreducible representations $[\alpha]$ of S_n belonging to a p-block characterized by a p-core of zero nodes namely

3.1 $$\sum_{\alpha} x_\alpha \sigma \lambda = 0,$$

where x_α denotes the degree of $[\alpha]$; $\sigma = (-1)^{\Sigma r_i} = \pm 1$ is the product of the parities of the p-hooks removable from the diagram $[\alpha]$ to yield the zero p-core; and λ is an integer $\geqslant 0$ which gives the multiplicity with which the star diagram $[\alpha]^*_p$ of $[\alpha]$ contains a chosen representation $[b]$ as an irreducible component.

For each choice of $[b]$ there arises an identity 3.1. In a recent paper by Todd (**12**) the same identities appear in another form, namely, as the expansions of the "new multiplication" of two Schur or S-functions of degrees m and n in terms of S-functions of degree mn, where the S-functions of degree n are the characters of irreducible representations of order n of the full linear group.

One can show that 3.1 actually admits of a more general interpretation with the degree of $[\alpha]$ replaced by its character χ_α, so that 3.1 *becomes an identity among the modular components of the irreducible representations.* Furthermore, these identities also exist for p-blocks characterized by non-zero p-cores.

Robinson's line of attack, however, does not yield the larger set of relations which arise from a consideration of the removal of just one hook from each of the Young diagrams of a given p-block, where this hook may be of length

$p, 2p, \ldots$, or bp. Suppose we start with the character relation

3.2 $$\sum_{\alpha} \chi_{\alpha}(R)\chi_{\alpha}(S) = 0,$$

where R and S do not belong to the same conjugate set of S_n. Let $R = V.P_k$, where P_k is a single cycle of length kp $(1 \leqslant k \leqslant b)$, and V is any permutation on the remaining $n - kp$ symbols. By the Murnaghan-Nakayama recursion formula

3.3 $$\chi_{\alpha}(V.P_k) = \sum_{\gamma_k} (-1)^{r_i}\chi_{\gamma_k}(V) = \sum_{\gamma_k} a_{\alpha\gamma_k}\chi_{\gamma_k}(V),$$

where the summation extends over all representations $[\gamma_k]$ of S_{n-kp} whose Young diagrams are obtainable from $[\alpha]$ by the removal of a single kp-hook H_i, and r_i is the leg length of H_i. Multiply 3.2 by $\chi_{\beta_k}(V)$, where $[\beta_k]$ is one of the irreducible representations of S_{n-kp} which appear in the right hand side of 3.3, and sum over all V:

$$\sum_{V} \chi_{\beta_k}(V) \sum_{\alpha} \chi_{\alpha}(S) \sum_{\gamma_k} a_{\alpha\gamma_k}\chi_{\gamma_k}(V) = 0.$$

Since the summation over V of the product $\chi_{\beta_k}(V) \cdot \chi_{\gamma_k}(V)$ yields zero in all cases except when $[\gamma_k] = [\beta_k]$, we obtain

$$\sum_{\alpha} \chi_{\alpha}(S) \sum_{V} \chi_{\beta_k}(V)a_{\alpha\beta_k}\chi_{\beta_k}(V) = 0.$$

This gives

3.4 $$\sum_{\alpha} a_{\alpha\beta_k}\chi_{\alpha}(S) = 0 \qquad (k = 1, 2, \ldots, b),$$

where $a_{\alpha\beta_k}$ is the parity of the kp-hook which is removed from $[\alpha]$ to yield $[\beta_k]$, and $[\beta_k]$ ranges over all diagrams of S_{n-kp} which appear as residual diagrams of $[\alpha]$. *Observe that the* $[\beta_k]$ *are those diagrams of* S_{n-kp} *with the same p-core as the original block of* $[\alpha]$'s. For each $[\beta_k]$, 3.4 is a linear relation among the characters χ_{α} of a fixed p-block which holds for all p-regular elements S of S_n, i.e., an identity among the modular components of these characters. A similar procedure applied to each of the other p-blocks yields further identities of the same type.

Example. The representations of S_8 which belong to the 2-block with 2-core [0] are [8], [7,1], [6,2], $[6,1^2]$, [5,3], $[5,1^3]$, $[4^2]$, [4,3,1], $[4,2^2]$, $[4,2,1^2]$, $[4,1^4]$, $[3^2,2]$ $[3^2,1^2]$, $[3,2^2,1]$, $[3,1^5]$, $[2^4]$, $[2^3,1^2]$, $[2^2,1^4]$, $[2,1^6]$, $[1^8]$. The necessary information for producing the identities appropriate to this 2-block is contained in the following table, in which the column labels are the various $[\beta_k]$ which appear after the removal of hooks of length $2k$ $(k = 1, 2, 3, 4)$ from the row labels $[\alpha]$, and the entries are the parities of these hooks:

[α]	k = 1										k = 2					k = 3		k = 4
	[6]	[5,1]	[4,2]	$[4,1^2]$	$[3^2]$	$[3,1^3]$	$[2^3]$	$[2^2,1^2]$	$[2,1^4]$	$[1^6]$	[4]	[3,1]	$[2^2]$	$[2,1^2]$	$[1^4]$	[2]	$[1^2]$	[0]
[8]	1	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1	.	1
[7,1]	.	1	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1	−1
[6,2]	1	.	1	.	.	.	.	.	.	.	.	.	1	.	.	.	−1	.
$[6,1^2]$	−1	.	.	1	.	.	.	.	.	.	.	.	.	1	.	.	.	1
[5,3]	.	1	.	.	1	.	.	.	.	.	.	.	−1	.	.	−1	.	.
$[5,1^3]$	.	−1	.	.	.	1	.	.	.	.	.	.	.	.	1	.	.	−1
$[4^2]$	.	.	1	.	−1	.	.	.	.	.	1	−1	.	.	.	.	.	.
[4,3,1]	.	.	.	1	.	.	.	.	.	.	−1	.	.	−1	.	1	.	.
$[4,2^2]$	.	.	1	−1	.	.	1	.	.	.	.	.	.	.	.	.	1	.
$[4,2,1^2]$	.	.	−1	.	.	.	.	1	.	.	1	.	.	.	−1	.	.	.
$[4,1^4]$	.	.	.	−1	.	.	.	.	1	.	−1	.	.	.	.	.	.	1
$[3^2,2]$	.	.	.	.	1	.	−1	.	.	.	.	−1	.	1	.	.	.	.
$[3^2,1^2]$	.	.	.	.	−1	1	.	−1	.	.	.	.	.	.	.	−1	.	.
$[3,2^2,1]$	.	.	.	.	.	−1	.	.	.	.	.	1	.	.	1	.	−1	.
$[3,1^5]$	.	.	.	.	.	−1	.	.	.	1	.	−1	.	.	.	.	.	−1
$[2^4]$	.	.	.	.	.	.	1	−1	.	.	.	.	.	1	−1	.	.	.
$[2^3,1^2]$	.	.	.	.	.	.	−1	.	−1	.	.	.	1	.	.	.	1	.
$[2^2,1^4]$	.	.	.	.	.	.	.	−1	.	−1	.	.	−1	.	.	1	.	.
$[2,1^6]$	.	.	.	.	.	.	.	.	−1	.	.	.	.	−1	.	−1	.	1
$[1^8]$	.	.	.	.	.	.	.	.	.	−1	.	.	.	.	−1	.	−1	−1

The columns of this table then give rise to the following identities:

$$
\begin{array}{lrlr}
k = 1: & \text{(i)} & [8] + [6,2] - [6,1^2] & = 0 \\
 & \text{(ii)} & [7,1] + [5,3] - [5,1^3] & = 0 \\
 & \text{(iii)} & [6,2] + [4^2] + [4,2^2] - [4,2,1^2] & = 0 \\
 & \text{(iv)} & [6,1^2] + [4,3,1] - [4,2^2] - [4,1^4] & = 0 \\
 & \text{(v)} & [5,3] - [4^2] + [3^2,2] - [3^2,1^2] & = 0 \\
 & \text{(vi)} & [5,1^3] + [3^2,1^2] - [3,2^2,1] - [3,1^5] & = 0 \\
 & \text{(vii)} & [4,2^2] - [3^2,2] + [2^4] - [2^3,1^2] & = 0 \\
 & \text{(viii)} & [4,2,1^2] - [3^2,1^2] - [2^4] - [2^2,1^4] & = 0 \\
 & \text{(ix)} & [4,1^4] - [2^3,1^2] - [2,1^6] & = 0 \\
 & \text{(x)} & [3,1^5] - [2^2,1^4] - [1^8] & = 0
\end{array}
$$

$$
\begin{array}{lrlr}
k = 2: & \text{(xi)} & [8] + [4^2] - [4,3,1] + [4,2,1^2] - [4,1^4] & = 0 \\
 & \text{(xii)} & [7,1] - [4^2] - [3^2,2] + [3,2^2,1] - [3,1^5] & = 0 \\
 & \text{(xiii)} & [6,2] - [5,3] + [2^3,1^2] - [2^2,1^4] & = 0 \\
 & \text{(xiv)} & [6,1^2] - [4,3,1] + [3^2,2] + [2^4] - [2,1^6] & = 0 \\
 & \text{(xv)} & [5,1^3] - [4,2,1^2] + [3,2^2,1] - [2^4] - [1^8] & = 0
\end{array}
$$

$$
\begin{array}{lrlr}
k = 3: & \text{(xvi)} & [8]-[5,3]+[4,3,1]-[3^2,1^2]+[2^2,1^4]-[2,1^6] & = 0 \\
 & \text{(xvii)} & [7,1] - [6,2] + [4,2^2] - [3,2^2,1] + [2^3,1^2] - [1^8] & = 0
\end{array}
$$

$$
\begin{array}{lrlr}
k = 4: & \text{(xviii)} & [8] - [7,1] + [6,1^2] - [5,1^3] + [4,1^4] & \\
 & & \qquad - [3,1^5] + [2,1^6] - [1^8] & = 0
\end{array}
$$

In general the identities that we have just derived will not be linearly independent. To establish their linear dependence, consider the character of a representation $[\alpha]$ for an element $R = P_u . P_v . W$, where P_u is a cycle of length up, P_v is a second cycle (distinct from P_u) of length vp $(u \neq v)$, and W is any permutation on the remaining $n - p(u + v)$ symbols. We assume that $[\alpha]$ contains hooks of length up and vp, and that $u + v \leqslant b$, where b is the number of successive p-hooks removable from $[\alpha]$ to yield its p-core. Applying the Murnaghan-Nakayama recursion formula twice, we obtain

$$
\begin{aligned}
\chi_\alpha(R) &= \sum_{\beta_u} a_{\alpha\beta_u} \chi_{\beta_u} (P_v . W) \\
&= \sum_{\beta_u} a_{\alpha\beta_u} \sum_{\beta_{u+v}} a'_{\beta_u\beta_{u+v}} \chi_{\beta_{u+v}}(W),
\end{aligned}
$$

if we think of removing a hook of length up first, and

$$
\begin{aligned}
\chi_\alpha(R) &= \sum_{\beta_v} a_{\alpha\beta_v} \chi_{\beta_v} (P_u . W) \\
&= \sum_{\beta_v} a_{\alpha\beta_v} \sum_{\beta_{u+v}} a''_{\beta_v\beta_{u+v}} \chi_{\beta_{u+v}} (W),
\end{aligned}
$$

if we remove a vp-hook first. Here $[\beta_u]$, $[\beta_v]$, $[\beta_{u+v}]$ are representations of S_{n-up}, S_{n-vp}, $S_{n-p(u+v)}$ respectively, and $a'_{\beta_u\beta_{u+v}}$ $(a''_{\beta_v\beta_{u+v}})$ is the parity of

the hook which must be removed from $[\beta_u]$ ($[\beta_v]$) in order to yield $[\beta_{u+v}]$. Since these are expressions for the same character, we have, for each appropriate $[\alpha]$ of the p-block,

3.5 $$\sum_{\beta_u} a_{\alpha\beta_u} \sum_{\beta_{u+v}} a'_{\beta_u\beta_{u+v}} \chi_{\beta_{u+v}}(W) = \sum_{\beta_v} a_{\alpha\beta_v} \sum_{\beta_{u+v}} a''_{\beta_v\beta_{u+v}} \chi_{\beta_{u+v}}(W),$$

a linear relation among the ordinary irreducible characters of $S_{n-p(u+v)}$ for all elements W of $S_{n-p(u+v)}$. The linear independence of these characters then implies

3.6 $$\sum_{\beta_u} a'_{\beta_u\beta_{u+v}} a_{\alpha\beta_u} = \sum_{\beta_v} a''_{\beta_v\beta_{u+v}} a_{\alpha\beta_v}$$

for each $[\beta_{u+v}]$. Observe that, if $u = v$, no relation of this kind arise, since 3.5 becomes simply an identity. Multiplying through 3.6 by $\chi_\alpha(S)$, where S is a p-regular element of S_n, and summing over the $[\alpha]$'s of the block under consideration, we obtain

$$\sum_{\beta_u} a'_{\beta_u\beta_{u+v}} \sum_{\alpha} a_{\alpha\beta_u} \chi_\alpha(S) = \sum_{\beta_v} a''_{\beta_v\beta_{u+v}} \sum_{\alpha} a_{\alpha\beta_v} \chi_\alpha(S).$$

For each $[\beta_{u+v}]$ this is a relation among the identities arising from the $[\beta_u]$'s and those arising from the $[\beta_v]$'s, where $u \neq v$.

Referring to our previous example, the only values of u and v ($u \neq v$) satisfying $u + v \leqslant b$, where $b = 4$ in this particular block, are 1,2 and 1,3, so that the number of relations among our identities is simply the number of $[\beta_3]$'s and $[\beta_4]$'s, namely 3. The relation arising from [2] is obtained by multiplying (i), (ii), . . . , (x) by $1, 0, 0, 0, -1, 0, 0, 1, -1, 0$ respectively (namely, the parities of the 4-hooks which must be removed from [6], [5,1], . . . , $[1^6]$ to yield [2], and 0 if no such 4-hook exists); (xi), (xii,) . . . , (xv) by $1, 0, 1, -1, 0$ respectively (namely, the parities of the 2-hooks removable from [4], [3,1], . . . , $[1^4]$ to yield [2]); and equating the two linear combinations to yield

$$\text{(i)} - \text{(v)} + \text{(viii)} - \text{(ix)} = \text{(xi)} + \text{(xiii)} - \text{(xiv)}.$$

Similarly, corresponding to $[1^2]$ and [0], we obtain:

$$\text{(ii)} - \text{(iii)} + \text{(vii)} - \text{(x)} = \text{(xii)} - \text{(xiii)} - \text{(xv)},$$

$$\text{(i)} - \text{(ii)} + \text{(iv)} - \text{(vi)} + \text{(ix)} - \text{(x)} = \text{(xvi)} - \text{(xvii)}.$$

We should not assume, however, that every relation among the identities which arises in this way is distinct from every other one; it may happen that one relation is simply a restatement of two or more other relations. Consider an element of the type $P_u \,.\, P_v \,.\, P_w \,.\, Q$, where the P's are defined as before and Q is any permutation on the remaining $n - p(u + v + w)$ symbols, $u \neq v \neq w$. Assuming that $u + v + w \leqslant b$, we obtain, by the same reasoning as before, the relations

$$\sum_{\beta_u} a_{s\beta_u} \sum_{\beta_{u+v+w}} a'_{\beta_u\beta_{u+v+w}} \chi_{\beta_{u+v+w}}(Q)$$
$$= \sum_{\beta_v} a_{s\beta_v} \sum_{\beta_{u+v+w}} a''_{\beta_v\beta_{u+v+w}} \chi_{\beta_{u+v+w}}(Q)$$
$$= \sum_{\beta_w} a_{s\beta_w} \sum_{\beta_{u+v+w}} a'''_{\beta_w\beta_{u+v+w}} \chi_{\beta_{u+v+w}}(Q),$$

and once again the linear independence of the characters $\chi_{\beta_{u+v+w}}$ of $S_{n-p(u+v+w)}$ yields

$$\sum_{\beta_u} a'_{\beta_u\beta_{u+v+w}} a_{s\beta_u} = \sum_{\beta_v} a''_{\beta_v\beta_{u+v+w}} a_{s\beta_v} = \sum_{\beta_w} a'''_{\beta_w\beta_{u+v+w}} a_{s\beta_w}$$

for each $[\beta_{u+v+w}]$. That is, for each $[\beta_{u+v+w}]$ only two of the three apparent relations which exist among the three sets of identities arising from the $[\beta_u]$'s, $[\beta_v]$'s, and $[\beta_w]$'s (namely, the relations between the sets taken in pairs) are distinct: the remaining relation is implied by the other two. The generalization of this to the case where we have $u_1, u_2, u_3, u_4, \ldots$, satisfying $u_1 + u_2 + u_3 + u_4 + \ldots \leqslant b$ and $u_1 \neq u_2 \neq u_3 \neq u_4 \ldots$, presents no added difficulty.

The following interpretation of the above relations among the identities may prove useful in understanding them. Since the number of modular irreducible characters of a group is less than the number of ordinary irreducible ones, there must exist a number of linear relations among the ordinary characters which hold for all p-regular elements (that is, identities among their modular components) in order to make up the difference. The number of modular characters of S_n being effectively the number of distinct partitions of n which contain neither p nor its multiples, the number of such identities must be the number of those partitions which contain p or its multiples, or the number of conjugate sets of p-singular elements. The identities that we have derived from all the blocks clearly correspond to those partitions of n which contain summands of length p, $2p$, $3p$, ... or bp, and the fact that the latter classification is not mutually exclusive (i.e. a partition may contain more than one multiple of p) means that we have more identities than there are partitions of this category. The relations among the identities serve to remove the duplications: however, their independence requires further study.

In our previous example, there were 16 p-singular conjugate sets and 19 identities(the modular characters of $[5,2,1]$ and $[3,2,1^3]$ of the block with 2-core $[3,2,1]$ satisfy the remaining identity $[5,2,1] - [3,2,1^3] = 0$), so that the three relations among the identities (namely the three relations corresponding to $[2]$, $[1^2]$, and $[0]$) make up the difference. The three relations may be accounted for by the conjugate sets $[6,2]$, $[4,2^2]$, $[4,2,1^2]$, which in a sense give rise to two identities each.

An examination of the number of identities in those p-blocks of S_n for which b is fixed leads to the following conjecture[1]:

[1]This has now been proved by G. de B. Robinson.

The number of indecomposables and the number of ordinary irreducible representations in a p-block of S_n characterized by a p-core of a nodes are the same as the corresponding numbers in a p-block of S_m with a p-core of a' nodes, where $n = a + bp$, $m = a' + bp$, i.e., where the same number of p-hooks are removable to yield a p-core. However, the corresponding blocks of the D-matrices will in general be different.

We conclude this section with two results which arise from the relations 3.4. Since these relations hold for all p-regular elements, we may replace the ordinary characters by their modular components Φ_λ and obtain

$$\sum_\alpha a_{\alpha\beta_k} \sum_\lambda d_{\alpha\lambda} \Phi_\lambda(S) = 0.$$

The linear independence of the Φ's then implies that

3.7 $$\sum_\alpha a_{\alpha\beta_k} d_{\alpha\lambda} = 0,$$

λ ranging over the modular characters of the block. Accordingly, if we think of the modular splitting of the $[\alpha]$'s as represented by a D-matrix (mod p) with the $[\alpha]$'s as row labels and the modular characters as column labels, we may state the following corollary to 3.7:

3.8. *The coefficients in the identities* 3.4 *are orthogonal to the columns of the D-matrix.*

Again, an indecomposable representation of the above block is a certain linear combination of ordinary irreducible representations, or, in terms of characters,

$$\eta_\lambda = \sum_\alpha d_{\alpha\lambda} \chi_\alpha.$$

Since the character χ_α for any element of the type $R = P_k \,.\, V$ takes the value $\sum_{\beta_k} a_{\alpha\beta_k} \chi_{\beta_k}(V)$, we have

$$\begin{aligned} \eta_\lambda(R) &= \sum_\alpha d_{\alpha\lambda} \sum_{\beta_k} a_{\alpha\beta_k} \chi_{\beta_k}(V) \\ &= \sum_{\beta_k} \left(\sum_\alpha a_{\alpha\beta_k} d_{\alpha\lambda}\right) \chi_{\beta_k}(V) \\ &= 0 \text{ by } 3.7. \end{aligned}$$

Since this holds for $k = 1, 2, \ldots, b$ and all p-singular elements of S_n are of the form $P_k \,.\, V$ for some integer k, we have a new proof of the following known result:

3.9. *The characters of the indecomposable representations of S_n vanish for all p-singular elements.*

4. The indecomposable representations of S_{2p}. We proceed with our investigation into the nature of the indecomposables for p-blocks of defects other

than 0 and 1 by applying Nakayama's induction formula

$$\eta^{(k)*} = \sum_{\lambda} a_{\kappa\lambda}\eta^{(\lambda)},$$

which operates on an indecomposable of a subgroup of S_n to produce a sum of indecomposables of S_n, to the particular case where the subgroup in question is S_{2p-1} and S_n is S_{2p}. This serves to effect a passage from p-blocks of defects 0 and 1 to p-blocks of higher defects, characterized by more than one p-hook in their Young diagrams.

It is necessary to consider only the p-block of S_{2p} characterized by a p-core of zero nodes, inasmuch as the theory is now complete for the one p-hook and the p-core cases. We shall verify that all the indecomposable representations of such a p-block may be obtained, via the inducing process, from the following two types of indecomposables of S_{2p-1}:

(i) the indecomposables of the p-blocks of S_{2p-1} with p-core $[p - r, 1^{r-1}]$ $r = 1, 2, \ldots, p-1$;

(ii) the indecomposable (and modular irreducible) representation $[p, 1^{p-1}]$.

The $p - 1$ indecomposables of the p-block of S_{2p-1} with p-core $[p - r, 1^{r-1}]$ are:

(1) $[2p - r, 1^{r-1}] + [p, p - r + 1, 1^{r-2}]$ $(r \neq 1)$

(1a) $[2p - 1] + [(p - 1)^2, 1]$ $(r = 1)$

(2) $[p - s, p - r + 1, 2^s, 1^{r-s-2}] + [p - s - 1, p - r + 1, 2^{s+1}, 1^{r-s-3}]$ $(s = 0, 1, 2, \ldots, r - 3)$

(3) $[p - r + 2, p - r + 1, 2^{r-2}] + [(p - r)^2, 2^{r-1}, 1]$

(4) $[p - r, p - r - t, 2^{r-1}, 1^{1+t}] + [p-r, p-r-t-1, 2^{r-1}, 1^{2+t}]$ $(t = 0, 1, 2, \ldots, p - r - 3)$

(5) $[p - r, 2^r, 1^{p-r-1}] + [p - r, 1^{p+r-1}]$ $(r \neq p - 1)$

(5a) $[3, 2^{p-2}] + [1^{2p-1}]$ $(r = p - 1)$

Neglecting all representations of S_{2p} except those belonging to the p-block with zero p-core, we can express the result of inducing on each of the above indecomposables of S_{2p-1} by the following notation:

(1)′ $[2p-r, 1^{r-1}] + [p, p-r+1, 1^{r-2}] \uparrow [2p-r+1, 1^{r-1}] + [2p-r, 1^r] + [p, p-r+2, 1^{r-2}] + [p, p-r+1, 1^{r-1}] = [a_1]+[b_1]+[c_1]+[d_1]$

(1a)′ $[2p-1]+[(p-1)^2,1] \uparrow [2p]+[2p-1, 1]+[p, p-1, 1]+[(p-1)^2,2]$

(2)′ $[p-s, p-r+1, 2^s, 1^{r-s-2}]+[p-s-1, p-r+1, 2^{s+1}, 1^{r-s-3}] \uparrow [p-s, p-r+2, 2^s, 1^{r-s-2}]+[p-s, p-r+1, 2^s, 1^{r-s-1}] + [p-s-1, p-r+2, 2^{s+1}, 1^{r-s-3}]+[p-s-1, p-r+1, 2^{s+1}, 1^{r-s-2}]$ $(s = 0, 1, 2, \ldots, r - 3)$

(3)′ $[p-r+2, p-r+1, 2^{r-2}]+[(p-r)^2, 2^{r-1}, 1] \uparrow [(p-r+2)^2, 2^{r-2}] +[p-r+2, p-r+1, 2^{r-2}, 1]+[p-r+1, p-r, 2^{r-1}, 1]+[(p-r)^2, 2^r]$

(4)′ $[p-r, p-r-t, 2^{r-1}, 1^{1+t}]+[p-r, p-r-t-1, 2^{r-1}, 1^{2+t}]$
$\uparrow [p-r+1, p-r-t, 2^{r-1}, 1^{1+t}]+[p-r, p-r-t, 2^{r}, 1^{t}]$
$+[p-r+1, p-r-t-1, 2^{r-1}; 1^{2+t}]+[p-r, p-r-t-1, 2^{r}, 1^{1+t}]$
$(t = 0, 1, 2, \ldots, p-r-3)$

(5)′ $[p-r, 2^{r}, 1^{p-r\,p-1}]+[p-r, 1^{p+r-1}] \uparrow [p-r+1, 2^{r}, 1^{p-r-1}]$
$+[p-r, 2^{r+1}, 1^{p-r-2}]+[p-r+1, 1^{p+r-1}]+[p-r, 1^{p+r}]$

(5a)′ $[3, 2^{p-2}]+[1^{2p-1}] \uparrow [3^2, 2^{p-3}]+[3, 2^{p-2}, 1]+[2, 1^{2p-2}]+[1^{2p}]$

The existence of an orthogonal relation 3.7 between the coefficients in the identities and the columns of the *D*-matrix suggests, as a first step towards building up our indecomposables from their irreducible components in any of the foregoing cases, the setting up of such a table as the following:

(1)′:

$$\begin{array}{c} \\ [a_1] \\ [b_1] \\ [c_1] \\ [d_1] \end{array}\begin{array}{c} \begin{array}{cccc} [p] & [p-r+1, 1^{r-1}] & [p-r, 1^{r}] & [0] \end{array} \\ \left[\begin{array}{cccc} . & 0 & . & r-1 \\ . & . & 0 & r \\ r-2 & 1 & . & . \\ r-1 & . & 1 & . \end{array}\right] \end{array}$$

For convenience in writing, the entries are not the parities $(-1)^{r_i}$, but the actual leg lengths r_i of the p-hooks and $2p$-hooks which are removable from the row labels to yield the column labels.

This differs from the table whose columns give us our identities for the p-block of representations of S_{2p} with p-core [0] in that the row labels of the latter table comprise all the representations of the p-block in question, while the table shown here contains only those representations of the p-block which arise from inducing on a single indecomposable of S_{2p-1}. Since 0 and the even integers (leg lengths) yield coefficients of $+1$ in the identities, and the odd integers coefficients of -1, it is clear that any linear combination of irreducible representations making up an indecomposable must contribute to each of these columns a number of odd integers equal to the number of even integers, if it contributes at all. The fact that the set of contributors to these columns is a sum of indecomposables means that the odd and even integers in each column do balance each other, so that the problem is to determine what further characterizations are needed to ensure that any subset of these contributors, possessing the same property, is actually an indecomposable.

Suppose we start by building the indecomposable to which $[a_1]$ belongs. The condition that we have imposed on our building process requires that $[c_1]$ be included in the combination in order that its 1 in column $[p-r+1, 1^{r-1}]$ balance the 0 of $[a_1]$. We say that these two p-hooks or, equivalently, the two entries which represent them, namely 0 and 1, are *linked* in column $[p-r+1, 1^{r-1}]$. Passing on to column [0], we observe that $[b_1]$ now must

be included in order that its r balance the $r - 1$ of $[a_1]$; likewise in column $[p]$ we observe that $[d_1]$ must be included in order that its $r - 1$ balance the $r - 2$ of $[c_1]$. At the same time the entries in $[p - r, 1^r]$ are linked. Since this brings into the fold all the representations at our disposal, the result is that the combination $[a_1] + [b_1] + [c_1] + d_1]$ is not a sum of indecomposables of S_{2p}, but an indecomposable by itself.

An indecomposable, then, appears to possess the property of having complete *linkages* in the columns of what we shall henceforth call its *linkage matrix*. In the case where only one p-hook can be removed from each of the Young diagrams of a p-block of S_n, such a matrix for an indecomposable of this p-block consists of only one column (headed by the p-core of the block) and two rows, and the entries are of the form r and $r + 1$ for $r = 0, 1, 2, \ldots, p - 2$. Hence the linkage matrix is a generalization of this latter case, where two representations of the block can be combined to form an indecomposable if and only if their p-hooks are *linked*, i.e., have *consecutive* leg lengths. The generalization lies in the presence of the linkage property in more than one column, where these additional columns stem from the fact that, when we deal with p-blocks of representations with Young diagrams containing more than one p-hook, more than one residual diagram arise after the removal of an initial p-hook; also the removal of hooks of length kp $(k > 1)$ needs to be considered. So far as satisfying a necessary condition for indecomposability is concerned (namely, that the character of an indecomposable representation vanishes for all p-singular elements), it would be sufficient that the number of even leg lengths balance the number of odd in each column of the linkage matrix; the only justification for our definition that linkage takes place, not haphazardly between hooks of odd and even leg lengths, but between hooks of *consecutive* leg lengths, is that, so far as the indecomposables of S_{2p} are concerned, our p-hooks and $2p$-hooks occur only in such pairs.

Linkage matrices (with leg lengths as entries) similar to that for (1)′ set out above show that likewise in each of the remaining cases all of the irreducible components must be taken in order to form a combination which is orthogonal to the identities.

Inducing on the indecomposable representation $[p, 1^{p-1}]$ in (ii) yields

$$[p + 1, 1^{p-1}] + [p, 2, 1^{p-2}] + [p, 1^p] = [a]+[b]+[c]$$

with linkage matrix;

$$\begin{array}{c} \\ {[a]} \\ {[b]} \\ {[c]} \end{array} \begin{array}{c} \begin{array}{ccc} [p] & [1^p] & [0] \end{array} \\ \begin{bmatrix} \cdot & 0 & p - 1 \\ p - 2 & 1 & \cdot \\ p - 1 & \cdot & p \end{bmatrix} \end{array}$$

so that $[a] + [b] + [c]$ clearly forms by itself an indecomposable of S_{2p}. Ob-

serve that, whereas each of the preceding indecomposables had four irreducible components, this last one is composed of only three.

Thus a single indecomposable of the p-block with zero p-core arises in each of these cases. There is, however, a certain amount of duplication as r ranges over its integral values. For instance, $r = 3$ gives the same indecomposable for case (2), $s = 0$, as $r = 1$ for case (4), $t = 0$; similarly $r = 4$ yields the same indecomposable for case (2), $s = 1$, as $r = 2$ for case (4) $t = 0$, and the same indecomposable for case (2), $s = 0$, as $r = 1$ for case (4), $t = 1$. The number of duplications (regarding an indecomposable as a duplication on its second appearance) is $0, 0, 1, 2, 3, \ldots, p - 3$, as r takes the values $1, 2, 3, 4, 5, \ldots, p - 1$ respectively, or $\frac{1}{2}(p - 2)(p - 3)$ in all. Hence the number of distinct indecomposables belonging to the designated p-block that arise in this way is

$$(p - 1)^2 + 1 - \tfrac{1}{2}(p - 2)(p - 3) = \tfrac{1}{2}(p - 1)(p + 2).$$

The ordinary representations of the p-block in question include the following:

$$\begin{aligned}
&[2p], [2p-1, 1], \ldots, [p+1, 1^{p-1}];\\
&[p^2], [p, p-1, 1], [p, p-2, 1^2], \ldots, [p, 1^p];\\
&[(p-1)^2, 2], [p-1, p-2, 2, 1], [p-1, p-3, 2, 1^2], \ldots,\\
&\qquad\qquad\qquad\qquad [p-1, 2^2, 1^{p-3}], [p-1, 1^{p+1}];\\
&[(p-2)^2, 2^2], [p-2, p-3, 2^2, 1], [p-2, p-4, 2^2, 1^2], \ldots,\\
&\qquad\qquad\qquad\qquad [p-2, 2^3, 1^{p-4}], [p-2, 1^{p+2}];\\
&\cdot \qquad \cdot \qquad \cdot \qquad \cdot \qquad \cdot \qquad \cdot\\
&[3^2, 2^{p-3}], [3, 2^{p-2}, 1], [3, 1^{2p-3}];\\
&[2^p], [2, 1^{2p-2}];\\
&[1^{2p}];
\end{aligned}$$

so that their number is $p + p + (p - 1) + (p - 2) + \ldots + 3 + 2 + 1 = p + \frac{1}{2}p(p + 1) = \frac{1}{2}p(p + 3)$. The number of identities that the characters of these representations must satisfy for all p-regular elements is simply the number of diagrams of S_p with zero p-core (i.e., the number of distinct p-hooks) together with the diagram of 0 nodes, or $p + 1$ in all. That these are linearly independent follows from the fact that there is no solution in positive integers of the inequality $u + v \leqslant 2$, $u \neq v$, since only these values of u and v can give rise to relations among the identities. Hence the number of modular irreducible representations or, equivalently, the number of indecomposables, is $\frac{1}{2}p(p + 3) - (p + 1) = \frac{1}{2}(p^2 + p - 2) = \frac{1}{2}(p - 1)(p + 2)$. A comparison of this number with the number obtained by the induction process shows that we have obtained all the indecomposables.

The following is the D-matrix for the 5-block of S_{10} with 5-core [0]. As noted before, each column contains exactly four 1's except the seventh, which contains three.

$$
\begin{array}{l}
[10] \\
[9,1] \\
[8,1^2] \\
[7,1^3] \\
[6,1^4] \\
[5^2] \\
[5,4,1] \\
[5,3,1^2] \\
[5,2,1^3] \\
[5,1^5] \\
[4^2,2] \\
[4,3,2,1] \\
[4,2^2,1^2] \\
[4,1^6] \\
[3^2,2^2] \\
[3,2^3,1] \\
[3,1^7] \\
[2^5] \\
[2,1^8] \\
[1^{10}]
\end{array}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

5. The indecomposable representations of S_n ($n > 2p$). In the section just concluded, the notion of a linkage has proved a useful tool in determining the indecomposables of S_{2p} belonging to the p-block with zero p-core, where a linkage was defined as taking place only between kp-hooks of consecutive leg lengths. Following this lead, we formulated a number of empirical rules regarding the use of linkages in constructing the indecomposables of S_n from those of S_{n-1}, and these rules produced (without apparent ambiguity) the indecomposables of S_n up to $n = 13$ for $p = 2, 3, 5$. The tables for most of these cases are contained in the author's thesis on file at the University of Toronto Library. A study of numerous examples led to a certain conjecture concerning the definition of an indecomposable and this will be the subject of a later paper.

References

1. R. Brauer, *On representations of groups of finite order*, Proc. Nat. Acad. Sc., vol. 25 (1939), 290-295.

2. ——— and C. Nesbitt, *On the modular character of groups*, Ann. of Math., vol. 42 (1941), 556-590.

3. ——— and G. de B. Robinson, *On a conjecture by Nakayama*, Trans. Roy. Soc. Can. (III), vol. 40 (1947), 11-25.

4. D. E. Littlewood and A. R. Richardson, *Group characters and algebras*, Phil. Trans. Roy. Soc. Lon. (A), vol. 233 (1934), 99-141.

5. T. Nakayama, *On some modular properties of irreducible representations of a symmetric group*, Part I, Jap. J. Math., vol. 17 (1940), 165-184.

6. ———, Part II, *ibid.*, vol. 17 (1941), 411-423.

7. G. de B. Robinson, *On the representations of the symmetric group*, Amer. J. Math., vol. 60 (1938), 745-760.

8. ———, Second Paper, *ibid.*, vol. 69 (1947), 286-298.

9. ———, Third Paper, *ibid.*, vol. 70 (1948), 277-294.

10. R. A. Staal, *Star diagrams and the symmetric group*, Can. J. Math., vol. 2 (1950), 79-92.

11. R. M. Thrall and C. J. Nesbitt, *On the modular representations of the symmetric group*, Ann. of Math., vol. 43 (1942), 656-670.

12. J. A. Todd, *A note on the algebra of S-functions*, Proc. Cambridge Phil. Soc., vol. 45 (1949), 328-334.

University of Toronto

ON RECURSIONS CONNECTED WITH SYMMETRIC GROUPS I

S. CHOWLA, I. N. HERSTEIN AND W. K. MOORE

ALTHOUGH the title of the paper suggests that the nature of the problem considered is group theoretic, our methods are almost completely combinatorial and number theoretic in nature, the group theory entering only insofar as it leads us to various recursions that we study. Let T_n denote the number of solutions of $x^2 = 1$ in S_n, the symmetric group of degree n. We proceed to find a recursion for T_n from which we obtain an explicit solution. From this we obtain an asymptotic value for T_n. We also exhibit some congruence and divisibility properties of the T_n. In a later paper we shall consider the problem of the number of solutions of $x^k = 1$ in S_n for k an arbitrary positive integer.

We begin with finding a recursion formula for the T_n, the number of solutions of $x^2 = 1$ in S_n. Although the derivation of this recursion is very simple, we give two proofs of it which, in a sense, are of a different mood. We assume $T_0 = T_1 = 1$.

LEMMA 1. $T_n = T_{n-1} + (n-1)T_{n-2}$.

First Proof. The only elements of order two in S_n are those which are the product of disjoint transpositions, and the unit element. The number of elements of order two which can be obtained from the permutations of the digits $1, 2, \ldots, n-1$, alone are T_{n-1}. The only other such elements are obtained from involving the digit n in a transposition with some other digit and multiplying by any other permutation of order two involving the remaining $n-2$ digits. Their number is clearly $(n-1)T_{n-2}$. Thus we obtain

$$T_n = T_{n-1} + (n-1)T_{n-2}. \tag{1}$$

Second Proof. It is well known that S_n is isomophic to the set of $n \times n$ matrices which have precisely one 1 in each row and column and zeros elsewhere. By direct checking it can be readily noted that the inverse of any such matrix is its transpose. So the question of the number of elements of order 2 in S_n becomes the question of finding how many self-adjoint matrices there are of the form described above. If the one in the top row occurs in the first column we are left an $(n-1) \times (n-1)$ matrix to consider, so the number of self-adjoint ones is T_{n-1}. If the one of the top row occurs in any other column, by the symmetry of the matrix, two rows and columns are used up, so we have an $(n-2) \times (n-2)$ matrix to consider, and we obtain T_{n-2}. Since the one in the top row could be put in $n-1$ such columns, the total number of this form is $(n-1)T_{n-2}$ and so again we obtain our recursion.

Received June 8, 1950; in revised form February 20, 1951.

From the recursion we obtain the following

LEMMA 2. $n^{\frac{1}{2}} \leqslant \dfrac{T_n}{T_{n-1}} \leqslant n^{\frac{1}{2}} + 1.$

Proof. The proof is by induction over n.

1. If $n = 1$, $T_1/T_0 = 1$, and the result is correct.

2. Suppose the result is correct for $n = r$. Consider T_{r+1}/T_r. Since

$$T_{r+1} = T_r + rT_{r-1},$$

$$T_{r+1}/T_r = 1 + r/(T_r/T_{r-1}) \leqslant 1 + r/r^{\frac{1}{2}} \leqslant 1 + (r+1)^{\frac{1}{2}}.$$

Also,

$$T_{r+1}/T_r = 1 + rT_{r-1}/T_r \geqslant 1 + r/(1 + r^{\frac{1}{2}}) \geqslant (r+1)^{\frac{1}{2}},$$

since

$$n = \left\{(n+1)^{\frac{1}{2}} - 1\right\}\left\{(n+1)^{\frac{1}{2}} + 1\right\} > \left\{(n+1)^{\frac{1}{2}} - 1\right\}(n^{\frac{1}{2}} + 1).$$

So the lemma follows from the induction.

From the lemma it follows trivially that:

THEOREM 3. *T_n/T_{n-1} is asymptotic to $n^{\frac{1}{2}}$.*

We again return to the recursion (1). Let $T_n = n!\,a_n$. Substituting in (1) we immediately obtain

$$na_n = a_{n-1} + a_{n-2}; \quad a_0 = a_1 = 1. \tag{2}$$

Consider the function $y = \sum_{i=0}^{\infty} a_i x^i$. We ask ourselves, what differential equation should y satisfy if the a_n satisfy the recursion in (2)? The differential equation suggested can be seen to be

$$x\,dy/dx = xy + x^2y.$$

Solving this by separating variables we see that

$$y = A \exp(x + \tfrac{1}{2}x^2).$$

Since $a_0 = 1$, $A = 1$. Thus we have:

THEOREM 4. *a_n is the coefficient of x^n in the power series expansion of* $\exp(x + \frac{1}{2}x^2)$.

Using the fact that

$$\exp(x + \tfrac{1}{2}x^2) = (\exp x)(\exp \tfrac{1}{2}x^2) = \sum_{i=0}^{\infty} \frac{x^i}{i!} \cdot \sum_{j=0}^{\infty} \frac{x^{2j}}{2^j j!} = \sum_{2j+i=n} \frac{x^n}{2^j j! i!},$$

we obtain

$$a_n = \sum_{2j+i=n} \frac{1}{2^j j! i!} \tag{3}$$

$$T_n = n! \sum_{2j+i=n} \frac{1}{2^j j! i!}. \tag{4}$$

On the other hand,

$$\exp(x + \tfrac{1}{2}x^2) = e^{-\frac{1}{2}} \exp\left(\frac{1+x}{2}\right)^2 = e^{-\frac{1}{2}} \sum_{n=0}^{\infty} \frac{(1+x)^{2n}}{2^n n!}\,.$$

whence,

$$a_{2m} = e^{-\frac{1}{2}} \frac{1}{2^m m!}\left(1 + \frac{2m+1}{2} + \frac{(2m+1)(2m+3)}{4} + \dots\right)$$

$$= e^{-\frac{1}{2}} \frac{1}{2^m m!} W_m = e^{-\frac{1}{2}} \frac{1}{2^m m!} (1 + \sum_{s=1}^{\infty} V_s),$$

where

$$V_s = \frac{(2m+1)(2m+3)(2m+5)\dots(2m+2s-1)}{(2s)!}\,.$$

Our first goal is an estimate of the size of W_m. This is given by

THEOREM 5. $W_m \sim \frac{1}{2}e^{\frac{1}{2}+(2m)^{\frac{1}{2}}}$.

To prove[1] Theorem 5, we need:

LEMMA 6 (Stirling's formula). *If x is a positive integer,*

$$\log x! = (x + \tfrac{1}{2}) \log x - x + \tfrac{1}{2} \log (2\pi) + O(x^{-1}).$$

LEMMA 7. *Suppose $b > a$ and that the interval (a, b) is divided into n equal parts of length h; let $f(x)$ be differentiable in (a, b) and $|f'(x)| \leqslant M$ in the interval. Then*

$$\left|\sum_{h=0}^{n-1} h\, f(a + rh) - \int_a^b f(x)dx\right| \leqslant h(b - a)\, M.$$

This lemma is an immediate consequence of the theory of Riemann integration and the first mean-value theorem of the differential calculus.

Consider V_s for

$$s = xm^{\frac{1}{2}} \tag{5}$$

and

$$\tfrac{1}{2} < x < 1 \qquad (m > m_0). \tag{6}$$

Now, using Lemma 6,

$$\log V_s = \sum_{t=1}^{s} \log(2m + 2t - 1) - \log(2s)!$$

$$= \sum_{t=1}^{s} \left\{\log 2m + \frac{2t-1}{2m} + O\left(\frac{t^2}{m^2}\right)\right\} - \{(2s + \tfrac{1}{2}) \log(2s) - 2s$$
$$+ \tfrac{1}{2} \log(2\pi) + O(s^{-1})\}$$

$$= s \log(2m) + \frac{s^2}{2m} + O\left(\frac{s^3}{m^2}\right)$$
$$- \{(2s + \tfrac{1}{2}) \log(2s) - 2s + \tfrac{1}{2} \log(2\pi) + O(m^{-\frac{1}{2}})\}$$

[1] We should like to thank Dr. W. R. Scott who carefully checked the proof of Theorem 5.

$$(7)\qquad = -xm^{\frac12}\log(2x^2) + 2xm^{\frac12} - \tfrac12\log(2xm^{\frac12}) - \tfrac12\log(2\pi) + \tfrac12 x^2 + O(m^{-\frac12}).$$

In (5), put

$$(8)\qquad x = 2^{-\frac12} + y$$

and restrict the limits of s by the inequality

$$(9)\qquad |y| \leqslant \epsilon = m^{-5/24}.$$

Since

$$(10)\qquad \log(1+t) = t - \tfrac12 t^2 + O(t^3),$$

for small t we obtain:

$$\log V_s = (2m)^{\frac12} + 2m^{\frac12}y - m^{\frac12}(2^{-\frac12}+y)\log(1 + 2^{3/2}y + 2y^2)$$
$$\qquad - \tfrac12\log\{(2m)^{\frac12} + 2m^{\frac12}y\} - \tfrac12\log(2\pi) + \tfrac14 + O(m^{-1/8})$$
$$= \tfrac14 - \tfrac12\log(2\pi) + (2m)^{\frac12} - \tfrac14\log(2m) - (2m)^{\frac12}y^2 + O(m^{\frac12}y^3) + O(m^{-1/8})$$
$$(11)\qquad = \tfrac14 - \tfrac12\log(2\pi) + (2m)^{\frac12} - \tfrac14\log(2m) - (2m)^{\frac12}y^2 + O(m^{-1/8}),$$

$$(12)\qquad V_s = \frac{e^{\frac14}}{(2\pi)^{\frac12}}\frac{e^{(2m)^{\frac12}}}{(2m)^{\frac14}}\, e^{-(2m)^{\frac12}y^2}\{1 + O(m^{-1/8})\}.$$

Hence

$$\sum_{\epsilon\leqslant y\leqslant\epsilon} V_s = \frac{e^{\frac14}}{(2\pi)^{\frac12}}\, e^{(2m)^{\frac12}} \sum_s \frac{e^{-(2m)^{\frac12}y^2}}{(2m)^{\frac14}}\,\{1 + O(m^{-1/8})\}$$

$$(13)\qquad = \frac{e^{\frac14}}{(4\pi)^{\frac12}}\, e^{(2m)^{\frac12}} \sum_s \left[\frac{2}{m}\right]^{\frac14} e^{-(2m)^{\frac12}y^2}\{1 + O(m^{-1/8})\},$$

where the summation is for all positive integers satisfying

$$|y| \leqslant \epsilon = m^{-5/24};$$

also $O(m^{-1/8})$ stands for $K_s m^{-1/8}$, where $|K_s| \leqslant |K$ an absolute constant for all s.

We proceed to show that

$$(14)\qquad \sum_s \left(\frac{2}{m}\right)^{\frac14} e^{-(2m)^{\frac12}y^2} \sim \int_{-\infty}^{\infty} e^{-w^2}\,dw = \pi^{\frac12}.$$

From (13) and (14) we finally obtain

$$(15)\qquad \sum_{|y|\leqslant\epsilon} V_s \sim \tfrac12 e^{\frac14} e^{(2m)^{\frac12}},$$

a result to which we shall return later.

To prove (14) we set

$$(16)\qquad y = \frac{w}{(2m)^{\frac14}}$$

and observe that as s increases by (1), x increases by $m^{-\frac{1}{2}}$, y increases by the same amount, and w increases by $\left(\frac{2}{m}\right)^{\frac{1}{2}}$. Since $w = w(s)$ is a function of s we can write

$$w(s+1) - w(s) = \left(\frac{2}{m}\right)^{\frac{1}{2}}, \tag{17}$$

the sum (14) becomes

$$\sum_{s_1 \leqslant s \leqslant s_2} e^{-w^2(s)}\{w(s+1) - w(s)\}, \tag{18}$$

where s_1 and s_2 are the smallest and largest positive integers s in the range $|y| \leqslant \epsilon$, i.e.

$$|w| \leqslant 2^{\frac{1}{2}}\, m^{1/24}. \tag{19}$$

From (17) and (19),

$$0 \leqslant w(s_1) + 2^{\frac{1}{2}} m^{1/24} \leqslant \left(\frac{2}{m}\right)^{\frac{1}{2}}, \tag{20}$$

and

$$0 \leqslant 2^{\frac{1}{2}} m^{1/24} - w(s_2) \leqslant \left(\frac{2}{m}\right)^{\frac{1}{2}}. \tag{21}$$

Clearly the sum in (18) is (by a crude estimate) equal to

$$\sum_{s_1 \leqslant s \leqslant s_2 - 1} e^{-w^2(s)}\{w(s+1) - w(s)\} + O(m^{-1/6}). \tag{22}$$

Now from Lemma 7,

$$\left| \sum_{s_1 \leqslant s \leqslant s_2 - 1} e^{-w^2}\left(\frac{2}{m}\right)^{\frac{1}{2}} - \int_{w_1}^{w_2} e^{-w^2}\, dw \right| = O\left(\frac{m^{1/24}}{m^{\frac{1}{4}}}\right) = O(m^{-5/24}), \tag{23}$$

where $w_1 = w(s_1)$, $w_2 = w(s_2)$. Clearly, $-w_1, w_2 \to \infty$. Also (14) follows from (17), (18), (19), (20), (21), (22), and (23). Hence we have established (15).

We next proceed to prove

$$\sum_{|y| > \epsilon} V_s = O\left(\frac{e^{(2m)^{\frac{1}{2}}}}{m^{1/24}}\right). \tag{24}$$

For $s \geqslant [(\frac{1}{2}m)^{\frac{1}{2}} + m^{\frac{1}{2}}\epsilon] = s_2$, $m > m_0$, we have

$$\begin{aligned} \frac{V_{s+1}}{V_s} = \frac{2m+2s+1}{(2s+1)(2s+2)} &\leqslant \frac{m}{2s^2} + O(m^{-\frac{1}{2}}) \\ &\leqslant \frac{1}{2(2^{-\frac{1}{2}} + \epsilon)^2} + O(m^{-\frac{1}{2}}) \\ &\leqslant 1 - 2^{\frac{1}{2}}\epsilon + O(\epsilon^2) + O(m^{-\frac{1}{2}}) \\ &\leqslant 1 - \epsilon, \end{aligned} \tag{25}$$

since $\epsilon = m^{-5/24}$. From (12),

$$V_{s_1} = O\left(\frac{e^{(2m)^{\frac{1}{2}}}}{m^{\frac{1}{4}}}\right). \tag{26}$$

So,

$$\sum_{s \geq s_1} V_s = O\left(\frac{e^{(2m)^{\frac{1}{2}}}}{m^{\frac{1}{4}}\epsilon}\right) = O\left(\frac{e^{(2m)^{\frac{1}{2}}}}{m^{1/24}}\right). \tag{27}$$

Similarly,

$$\sum_{s < s_1} V_s = O\left(\frac{e^{(2m)^{\frac{1}{2}}}}{m^{1/24}}\right). \tag{28}$$

From (15), (27), and (28),

$$W_m = 1 + \sum_{s=1}^{\infty} V_s \sim \tfrac{1}{2}e^{\frac{1}{4}+(2m)^{\frac{1}{2}}}, \tag{29}$$

establishing Theorem 5.

Now for $n = 2m$,

$$T_n = n!\,\frac{e^{-\frac{1}{4}}}{2^m m!}\,W_m \sim \left(\frac{n}{e}\right)^{\frac{1}{2}n}\,\frac{e^{n^{\frac{1}{2}}}}{2^{\frac{1}{2}}e^{\frac{1}{4}}}. \tag{30}$$

For odd n, $n = 2m + 1$,

$$T_{2m+1} \sim T_{2m}(2m)^{\frac{1}{2}},$$

from Theorem 3. Thus

$$T_{2m+1} \sim 2^{-\frac{1}{2}}e^{-\frac{1}{4}}\left(\frac{2m}{e}\right)^{m} e^{(2m)^{\frac{1}{2}}}(2m)^{\frac{1}{2}}. \tag{31}$$

But

$$\frac{\left(\frac{2m+1}{e}\right)^{m}\left(\frac{2m+1}{e}\right)^{\frac{1}{2}}}{(2m)^{\frac{1}{2}}\left(\frac{2m}{e}\right)^{m}}\;\frac{e^{(2m+1)^{\frac{1}{2}}}}{e^{(2m)^{\frac{1}{2}}}} \to 1 \tag{32}$$

as $m \to \infty$. So

$$T_{2m+1} \sim \frac{1}{2^{\frac{1}{2}}}\,\frac{1}{e^{\frac{1}{4}}}\left(\frac{2m+1}{e}\right)^{\frac{1}{2}(2m+1)} e^{(2m+1)^{\frac{1}{2}}}, \tag{33}$$

and (30) and (33) together prove

THEOREM 8. $$T_n \sim \frac{(n/e)^{\frac{1}{2}n}e^{n^{\frac{1}{2}}}}{2^{\frac{1}{2}}e^{\frac{1}{4}}}.$$

We now turn to some other properties of the T_n's. These results on divisibility and congruences of the T_n's, while they are very easy to prove, are of some interest.

We first prove

THEOREM 9. *If m is an odd integer, then $T_{n+m} \equiv T_n \pmod{m}$.*

Proof. It is clear that it is sufficient to prove the theorem for prime powers. The proof for these is exactly the same as the proof for odd primes. So we prove the theorem for odd primes. The proof is by induction over n, where $m = p$, a prime.

(i) If $n = 0$, then $T_p = p! \sum_{2i+j=n} \frac{1}{2^i i! j!}$ and this is clearly congruent to $1 = T_0$ modulo p, if p is an odd prime.

(ii) If $n = 1$, $T_{p+1} = T_p + (p + 1 - 1)T_{p-1}$ and this is congruent to T_p modulo p, hence to 1; that is $T_{p+1} \equiv T_1 \pmod{p}$.

(iii) Suppose that $T_{r+p} \equiv T_r \pmod{p}$. Now

$$T_{r+1+p} = T_{r+p} + (r + p)T_{r+p-1} \equiv T_r + rT_{r-1} \pmod{p},$$

by our induction. Since $T_{r+1} = T_r + rT_{r-1}$, our result follows.

The other number theoretic property of T_n that we prove is that it is highly divisible by powers of 2. In fact, we prove

THEOREM 10. *If $n \geqslant 4s - 2$, then 2^s divides T_n.*

Proof. By induction over s.

(i) If $s = 1$, since all the T_n's for $n \geqslant 2$ are even (as can be easily seen from the recursion), the result is correct.

(ii) Suppose that if $n \geqslant 4r - 2$, $2^r | T_n$. Let $n \geqslant 4(r + 1) - 2 = 4r + 2$.

$$\begin{aligned} T_n = T_{n-1} + (n-1)T_{n-2} &= nT_{n-2} + (n-2)T_{n-3} \\ &= (2n-2)T_{n-3} + n(n-3)T_{n-4}. \end{aligned}$$

Since $n - 4 \geqslant 4r - 2$, then $2^r | T_{n-3}$ and $2^r | T_{n-4}$. Since the coefficients of each of these in the expression for T_n is even, $2^{r+1} | T_n$. This concludes the induction and proves the theorem.

We should like to make one remark. In Theorem 3 we proved that $T_n/T_{n-1} \sim n^{\frac{1}{2}}$. We feel that much more is true, namely that $T_n/T_{n-1} = n^{\frac{1}{2}} + A + Bn^{-\frac{1}{2}} + Cn^{-1} + Dn^{-3/2} + \ldots$, for appropriate constants $A, B, C, D, \ldots$.

University of Kansas

ON THE NATURE OF THE SPECTRUM OF SINGULAR SECOND ORDER LINEAR DIFFERENTIAL EQUATIONS

E. A. CODDINGTON AND N. LEVINSON

LET $p(x) > 0$, $q(x)$ be two real-valued continuous functions on $0 \leqslant x < \infty$. Suppose that the differential equation with the real parameter λ

$$(1) \qquad (py')' + (\lambda - q)y = 0$$

does not possess two linearly independent solutions of class $L^2(0, \infty)$ for some λ. According to the Weyl classification [6] equation (1) is then said to be of the *limit-point* type. In this case (1) together with a boundary condition

$$(2) \qquad y(0) \cos \alpha + y'(0) \sin \alpha = 0$$

determine an eigenvalue problem.

For every α in (2) there corresponds a spectrum $S(\alpha)$ which includes a (possibly empty) point spectrum $P(\alpha)$, the latter set consisting of those λ for which there exists a real solution of (1) satisfying (2) and is of class $L^2(0, \infty)$. The derivative $S'(\alpha)$ of $S(\alpha)$ consists of the continuous spectrum and the cluster points of the point spectrum. Using the theory of bounded quadratic forms whose differences are completely continuous, and the idea of a Hellinger integral, Weyl proved [5, p. 378; 6, p. 251] that the set $S'(\alpha)$ does not depend on α, and hence can be denoted simply by S'.

Recently one of the authors [3] gave a simplified proof of the Parseval relation corresponding to the problem (1) and (2). This relation was obtained by a natural limiting process on the Parseval equality which holds for the corresponding Sturm-Liouville problem on a bounded interval $0 \leqslant x \leqslant b < \infty$. It turns out that the formulae (but not the end result) developed in this proof (which were obtained by Titchmarsh [4] using other methods[1]) can be used to obtain a direct proof of the invariance of the set S'. Also the oscillation and separation theorems due to Hartman and Wintner [1], [2] can be obtained in a similar way (see (II), (III) below).

Denote by $S^*(\alpha)$ the complement of the set $S'(\alpha)$ with respect to the set $-\infty < \lambda < \infty$. Since $S'(\alpha)$ is closed, $S^*(\alpha)$ is open. In this notation we prove:

(I) *The set of cluster points $S'(\alpha)$ of the spectrum $S(\alpha)$ for the problem* (1) *and* (2) *is independent of α, and hence can be denoted by S'.*

Received March 24, 1950.

[1]K. Kodaira, *American Journal of Mathematics*, vol. 71 (1949), pp. 921-945, obtains Titchmarsh's results and also a formula for $m(\lambda)$ using the spectral representation of a self-adjoint operator in Hilbert space.

(II) *If* $\lambda \in S^*$, *(S^* being the complement of S') then* $\lambda \in P(\alpha)$ *for some* $\alpha = \alpha(\lambda)$.

(III) *The function* $\alpha(\lambda)$ *for which* $\lambda \in P(\alpha)$ *is regular, monotone increasing on* S^*.

Proof of I. We need some known facts. Let $\theta(x, \lambda)$ and $\phi(x, \lambda)$ be solutions of (1) which satisfy

$$p(0)\theta(0, \lambda) = \cos \alpha, \quad p(0)\theta'(0, \lambda) = \sin \alpha \tag{3}$$
$$\phi(0, \lambda) = \sin \alpha, \quad \phi'(0, \lambda) = -\cos \alpha.$$

For $\lambda = u + iv$, $v \neq 0$, consider the solution of (1):

$$\psi_b(x, \lambda) = \theta(x, \lambda) + l_b(\lambda)\, \phi(x, \lambda)$$

which satisfies a real boundary condition at $x = b$,

$$\psi_b(b, \lambda) \cos \beta + \psi'_b(b, \lambda) \sin \beta = 0.$$

For each b, as β varies, $l_b(\lambda)$ describes a circle C_b in the complex plane, and it can be shown that as $b \to \infty$, the circles C_b converge either to a limit-circle or to a limit-point. Since we are assuming the limit-point case, let us denote this point by $m(\lambda) = m(\lambda, \alpha)$. For any $v \neq 0$ it is true that, if

$$\psi(x, \lambda) = \theta(x, \lambda) + m(\lambda)\, \phi(x, \lambda), \tag{4}$$

then

$$\int_0^\infty |\psi(x, \lambda)|^2 dx \leqslant -\frac{\Im(m)}{v}. \tag{5}$$

Also the function $m(\lambda)$ is analytic in either half-plane $v > 0$, $v < 0$. It was shown in [3] that there exists a monotone non-decreasing function $\rho(\sigma) = \rho(\sigma, \alpha)$ on $-\infty < \sigma < \infty$ and a real constant c such that

$$\frac{-\Im(m)}{v} = \int_{-\infty}^\infty \frac{d\rho(\sigma)}{(u - \sigma)^2 + v^2} + c. \tag{6}$$

It follows from the proof of (6) that c is a non-negative constant, and as a matter of fact this can be easily seen from (6) and (5) by letting $v \to \infty$.

An immediate consequence of (6) is that at points of continuity $\sigma = \sigma_1$, $\sigma = \sigma_2$ of $\rho(\sigma)$

$$\rho(\sigma_2) - \rho(\sigma_1) = \lim_{v \to 0} \frac{1}{\pi} \int_{\sigma_1}^{\sigma_2} -\Im(m(u + iv))du. \tag{7}$$

In proving (6) it was of course shown that the integral

$$\int_{-\infty}^\infty \frac{d\rho(\sigma)}{1 + \sigma^2} \tag{8}$$

was convergent. If one considers the function of λ given by

$$\int_{-\infty}^\infty \left[\frac{1}{\lambda - \sigma} + \frac{1}{\sigma + i}\right] d\rho(\sigma) - c\lambda + d$$

where $c > 0$ is the constant in (6) and d is a (complex) constant, then the integral obviously exists by virtue of the convergence of (8). Also the imaginary part of this function is identical with $\mathfrak{I}(m)$ (see (6)), if

$$\mathfrak{I}(d) = \int_{-\infty}^{\infty} \frac{d\rho(\sigma)}{1 + \sigma^2}.$$

Therefore, since $m(\lambda)$ and this function are regular on $v > 0$, the latter coincides with $m(\lambda)$ except for a real constant, which may be incorporated into d. Hence

$$(9) \qquad m(\lambda) = \int_{-\infty}^{\infty} \left[\frac{1}{\lambda - \sigma} + \frac{1}{\sigma + i} \right] d\rho(\sigma) - c\lambda + d.$$

For a given α, the spectrum $S(\alpha)$ is the σ set which is the complement of the set of points in the neighbourhood of which $\rho(\sigma, \alpha)$ is constant. The jumps of $\rho(\sigma, \alpha)$ correspond to the point spectrum $P(\alpha)$. Clearly $\rho(\sigma, \alpha)$, considered as a function on $S^*(\alpha)$, is constant except for isolated jumps. Thus $m(\lambda, \alpha)$ by (9) is analytic on $S^*(\alpha)$ except at the isolated discontinuities of $\rho(\sigma, \alpha)$ where $m(\lambda, \alpha)$ has simple poles. Also $m(u, \alpha)$ is real on $S^*(\alpha)$.

Consider the boundary condition (2) corresponding to α_1, α_2 where $\alpha_1 \not\equiv \alpha_2$ (mod π), and define the solutions $\theta(x, \lambda, \alpha_i)$, $\phi(x, \lambda, \alpha_i)$, $\psi(x, \lambda, \alpha_i)$, $(i = 1, 2)$ by (3) and (4). From these relations it is clear that if $\gamma = \alpha_2 - \alpha_1$, then

$$\begin{aligned} \phi(x, \lambda, \alpha_1) &= \cos \gamma\, \phi(x, \lambda, \alpha_2) - p(0) \sin \gamma\, \theta(x, \lambda, \alpha_2) \\ p(0)\theta(x, \lambda, \alpha_1) &= \sin \gamma\, \phi(x, \lambda, \alpha_2) + p(0) \cos \gamma\, \theta(x, \lambda, \alpha_2). \end{aligned}$$

Consequently

$$p(0)\psi(x, \lambda, \alpha_1) =$$
$$[\cos\gamma - p(0)m(\lambda, \alpha_1)\sin\gamma]p(0)\,\theta(x, \lambda, \alpha_2) + [\sin\gamma + p(0)m(\lambda, \alpha_1)\cos\gamma]\phi(x, \lambda, \alpha_2),$$

and since $\psi(x, \lambda, \alpha_1)$, $\psi(x, \lambda, \alpha_2)$ are both of class $L^2(0, \infty)$ one is a constant multiple of the other by virtue of the limit-point assumption. This implies that[2]

$$(10) \qquad p(0)m(\lambda, \alpha_2) = \frac{\sin \gamma + p(0)m(\lambda, \alpha_1) \cos \gamma}{\cos \gamma - p(0)m(\lambda, \alpha_1) \sin \gamma}, \gamma = \alpha_2 - \alpha_1.$$

Since $m(\lambda, \alpha_1)$ is meromorphic on $S^*(\alpha_1)$, so is $m(\lambda, \alpha_2)$ and because $m(u, \alpha_1)$ is real on $S^*(\alpha_1)$, it follows that for values of u on this set $\mathfrak{I}(m(\lambda, \alpha_2)) \to 0$, $v \to +0$, except at isolated poles of $m(\lambda, \alpha_2)$. From this fact we see from (7) that $\rho(\sigma, \alpha_2)$ is constant on $S^*(\alpha_1)$ except for jumps at isolated poles of $m(\lambda, \alpha_2)$. This proves that $S^*(\alpha_2) \supset S^*(\alpha_1)$. Since the roles of α_1 and α_2 can be interchanged we get $S^*(\alpha_2) = S^*(\alpha_1) = S^*$, and hence $S'(\alpha_2) = S'(\alpha_1)$, thus proving (I).

[2]The boundary condition (2) is often replaced by $y(0) \cos \alpha + p(0)y'(0) \sin \alpha = 0$, which has the effect of eliminating $p(0)$ in (10).

Proof of (II) *and* (III). From (9) it is clear that, except for poles of $m(\lambda, \alpha_1)$, on S^*

$$\frac{\partial m(u, \alpha_1)}{\partial u} = -\int_{-\infty}^{\infty} \frac{d\rho(\sigma)}{(u - \sigma)^2} - c$$

and hence $\dfrac{\partial m(u, \alpha_1)}{\partial u} < 0$. We see therefore that on S^*, $m(u, \alpha_1)$ is a regular monotone decreasing function of u, except for poles of $m(\lambda, \alpha_1)$.

Let $\lambda_1 \in P(\alpha_1)$ and suppose $\lambda \in S^*$ ($= S^*(\alpha_1)$) is in a sufficiently small open interval about λ_1 which contains no other points of $P(\alpha_1)$. Then it follows from (10) that $m(\lambda, \alpha_2)$ will have a pole for some $\alpha_2 \not\equiv \alpha_1 \pmod{\pi}$ determined by the relation $p(0)m(\lambda, \alpha_1) = \cot \gamma$, that is, there exists an $\alpha_2 = \alpha_2(\lambda)$ for which $\lambda \in P(\alpha_2)$. Every $\lambda \in S^*$ can be brought within an open interval containing at most one point of $P(\alpha_1)$, so that $\alpha_2 = \alpha_2(\lambda)$ exists for all $\lambda \in S^*$, even at the points λ_1 where $\alpha_2 \equiv \alpha_1 (\text{mod } \pi)$ and $\lambda_1 \in P(\alpha_1)$. This proves (II).

From the relation $p(0)m(\lambda, \alpha_1) = \cot \gamma = \cot(\alpha_2 - \alpha_1)$ it follows that

$$\alpha_2(\lambda) = \alpha_1 + \cot^{-1}[p(0)m(\lambda, \alpha_1)]. \tag{11}$$

Moreover, from the discussion above, since $m(\lambda, \alpha_1)$ is regular on an interval about λ_1, $\alpha_2(\lambda)$ is a regular function of λ on such an interval and

$$\frac{d\alpha_2(\lambda)}{d\lambda} = \frac{-p(0)}{[1 + (p(0)m(\lambda, \alpha_1))^2]} \frac{\partial m(\lambda, \alpha_1)}{\partial \lambda}. \tag{12}$$

But $p(0) > 0$, $\dfrac{\partial m(\lambda, \alpha_1)}{\partial \lambda} < 0$ on S^* (except for poles of $m(\lambda, \alpha_1)$, and therefore $\dfrac{d\alpha_2(\lambda)}{d\lambda} > 0$ on an open interval about λ_1, that is, $\alpha_2(\lambda)$ is increasing on S^*. This completes the proof of (III).

References

[1] P. Hartman and A. Wintner, *An oscillation theorem for continuous spectra*, Proc. Nat. Acad. Sci., vol. 33 (1947), 376-379.

[2] ———, *A separation theorem for continuous spectra*, Amer. J. Math., vol. 71 (1949), 650-662.

[3] N. Levinson, *A simplified proof of the expansion theorem for singular second order linear differential equations*, Duke Math. J. vol. 18 (1951), 57-71.

[4] E. C. Titchmarsh, *Eigenfunction expansions associated with second-order differential equations*, (Oxford, 1946).

[5] H. Weyl, *Ueber beschränkte quadratische Formen, deren Differenz vollstetig ist*, Rend. Circ. Palermo, vol. 27 (1909), 373-392.

[6] H. Weyl, *Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen*, Math. Ann., vol. 68 (1910), 222-269.

Massachusetts Institute of Technology

SUR LES SOLUTIONS PÉRIODIQUES DE CERTAINS SYSTÈMES DIFFÉRENTIELS PERTURBÉS

GEORGES REEB

Introduction. Les systèmes différentiels dont il est question dans ce travail, sont du type suivant:

$$(1) \qquad dy = E_\mu(y)dt. \qquad (\Sigma_\mu)$$

E_μ est un champ de vecteurs défini sur une variété V_{n+1} à $n+1$ dimensions, trois fois continûment différentiable, jouissant des propriétés suivantes:

(a) *E_μ dépend du paramètre réel et positif μ ($\mu \geqslant 0$).*

(b) *Le vecteur $E_\mu(y)$ du champ E_μ, attaché au point y de V_{n+1}, est une fonction deux fois continûment différentiable du couple (y, μ).*

(c) *$E_0(y) \neq 0$ pour tout $y \in V_{n+1}$.*

(d) *Les trajectoires du champ E_0 sont fermées (et par suite homémorphes au cercle S_1). De plus elles sont les fibres d'une structure fibrée de V_{n+1}, dont la base est une variété V_n à n dimensions. L'application canonique P de V_{n+1} sur V_n est alors deux fois continûment différentiable.*

En fait les hypothèses de différentiabilité ne seront pas systématiquement explicitées dans la suite. On pourra par exemple se contenter de supposer que toutes les fonctions, variétés, etc... qui interviennent sont indéfiniment continûment différentiables.

On se propose d'étudier, *par des méthodes simples et pour la plupart classiques*, certaines propriétés des trajectoires fermées du champ E_μ pour les petites valeurs de μ. [**16**; **19**; **18**; **28**].

En faisant les hypothèses (a), (b), (c) et (d) sur le champ E_μ on peut étendre au cas d'un système à plusieurs degrés de liberté un certain nombre des résultats connus [**19**, p. 184-204] pour l'équation différentielle:

$$(2) \qquad x'' + x = \mu f(x,x'),$$

ou plus généralement

$$(3) \qquad x'' + g(x) = \mu f(x,x').$$

Les équations (2) et (3) comprennent comme cas particulier l'équation différentielle des oscillations de relaxation. (On remarquera que si dans (2) on donne au paramètre μ la valeur 0 ce système admet des solutions périodiques, et par conséquent vérifie les hypothèses (a), (b), (c), (d) (cf. §**2.1**).

Le chapitre II est consacré à l'étude de certains champs E_0 (possédant la

Reçu le 13 octobre, 1949; revisé le 21 février, 1951.

propriété (d)) qui se présentent souvent dans des problèmes de dynamique. Dans le cas particulier où il existe un invariant intégral du type de la dynamique, on remarque que la période est la même pour toutes les trajectoires correspondant à une valeur donnée de l'hamiltonien H. D'autres propriétés globales résultent encore de l'existence de l'invariant intégral.

Dans le chapitre III on associe au champ E_μ un champ $\bar{E}$, défini sur la base V_n, et on étudie les relations entre les trajectoires de $\bar{E}$ et E_μ. Il s'agit de façon plus précise de relations entre les singularités de $\bar{E}$ et les trajectoires fermées de E_μ. L'idée essentielle peut au fond se résumer ainsi: les propriétés qualitatives des trajectoires du champ E_μ sont analogues à celles des trajectoires obtenues en se contentant d'une première approximation, [**18**]. C'est la méthode des petits paramètres de H. Poincaré.

Dans IV on suppose que le champ E_μ admet l'invariant intégral de M. Élie Cartan [**7**], et on en déduit qu'il en est de même du champ $\bar{E}$. Les théorèmes de M. Morse [**21**] permettent alors de préciser les résultats de III [**28**].

Le dernier chapitre VI comprend quelques remarques sur l'existence de solutions périodiques de certains systèmes (Σ_μ), sous des conditions plus larges que celles des chapitres précédents.

Il a semblé utile de résumer en I certaines définitions et propriétés fréquemment utilisées par la suite.

I. Rappel de certaines propriétés et définitions

1.1. Formes différentielles extérieures (notations). Soit V_n une variété numérique. Une forme différentielle extérieure sur V_n est désignée par une lettre grecque α, ω, π, . . ; il n'est pas nécessaire de mettre en évidence le degré de la forme considérée. La loi de produit extérieur de formes extérieures est notée $\otimes$. On utilise la notation $(\alpha)^2$ pour $\alpha \otimes \alpha$, $(\alpha)^p$ pour $\alpha \otimes \alpha \otimes \ldots \otimes \alpha$, ($p$ facteurs). La différentielle extérieure d'une forme extérieure π est notée $d\pi$. L'opérateur d augmente les degrés d'une unité et jouit des propriétés suivantes

$$d(\alpha + \beta) = d\alpha + d\beta,\ dd\alpha = 0, \quad d(\lambda\alpha) = \lambda d\alpha + d\lambda \otimes \alpha$$

(où λ est une fonction numérique), etc. On dit que π est fermé si $d\pi = 0$; on dit que π est homologue à zéro, s'il existe α tel que $\pi = d\alpha$. Si V_n est compact et si α est une forme fermée de degré n qui ne s'annule en aucun point de V_n, la forme α n'est pas homologue à zéro. En effet $\int_{V'_n} \alpha \neq 0$ sur une composante connexe V'_n de V_n On pourra se reporter aux ouvrages suivants: [**2**, Livre II, Chapitre III; **6**, p. 1-32] algèbre extérieure; [**6**, p. 33-44] différentation extérieure; [**3**] anneau de cohomologie de V_n. Voir aussi [**9**, p. 146-152; **7**].

1.2. Invariants intégraux et systèmes différentiels. Soit (Σ) un système différentiel ordinaire sur V_n. Si x_i $(i = 1, \ldots, n)$ sont des coordonnées locales sur V_n, le système (Σ) peut se mettre sous la forme:

$$\frac{dx_1}{X_1} = \frac{dx_2}{X_2} = \ldots = \frac{dx_n}{X_n}.$$

Le système (Σ) définit un champ de vecteurs E ayant pour composantes X_i. D'une façon précise E est défini à un facteur numérique près; en d'autres termes le système (Σ) définit un champ de directions sur V_n. Réciproquement un champ de directions ou un champ de vecteurs E sur V_n définit un système différentiel (Σ). Soit $\phi_1, \ldots, \phi_{n-1}$, un système complet d'intégales premières de (Σ)

DÉFINITION (a). *La forme différentielle extérieure π est un invariant intégral absolu du système différentiel (Σ) (ou du champ E) si π est une forme différentielle construite sur les différentielles $d\phi_i (i=1, \ldots, n-1)$ dont les coefficients sont des fonctions de $\phi_1, \ldots, \phi_{n-1}$. Un invariant intégral relatif est une forme π dont la différentielle extérieure $d\pi$ est un intégral absolu.*

La propriété suivante est une conséquence immédiate de la définition (a).

THÉORÈME (a). *Soient $\phi_1, \ldots, \phi_{n-1}, \phi_n$, des coordonnées locales dans V_n, soient $\phi_1, \ldots, \phi_{n-1}$ des intégrales premières de (Σ). La transformation F: $\phi_i \to \phi_i$ ($i = 1, \ldots, n-1$), $\phi_n \to f(\phi_1, \ldots, \phi_n)$, laisse invariant tout invariant intégral absolu π de (Σ). En d'autres termes $F^*(\pi) = \pi$, où F^* est l'application transposée de F.*

Le théorème (a) permet de démontrer:

THÉORÈME (a'). *Soit F une application d'un ouvert Ω de V_n dans V_n, qui applique tout point x de Ω sur un point la trajectoire de (Σ) issue de x. Si π est un invariant intégral absolu de (Σ), alors $F^*(\pi) = \pi$.*

Soit π une forme différentielle extérieure fermée de degré deux sur V_n.

DÉFINITION (b). *On appelle classe de la forme fermée de degré deux π, le nombre entier q tel que $(\pi)^q \neq 0$ en tout point et $(\pi)^{q+1} = 0$.*

La classe de π est évidemment inférieure ou égale à $\frac{1}{2}n$.

THÉORÈME (b). *Si n est pair ($n = 2q$) et si π est fermé et de classe q, on peut trouver des coordonnées locales $p_1, q_1, p_2, q_2, \ldots, p_q, q_q$, telles que $\pi = \sum_{i=1}^{q} dp_i \otimes dq_i$; de plus on peut prendre pour p_1 une fonction arbitraire* [**6**, p. 52; **7**, p. 29; **5** p. 121-134].

THÉORÈME (c). *Soient p_i, q_i ($i = 1, \ldots, r$) et t des coordonnées locales dans V_n ($n = 2r + 1$). La forme extérieure $\pi = \Sigma_i dp_i \otimes dq_i - dH \otimes dt$, où H est une fonction numérique sur V_n, est fermée et est de classe r. Il existe exactement un champ de directions E pour lequel π est un invariant intégral absolu. Le système différentiel lié à E est*

$$(\Sigma) \qquad \frac{dp_i}{\partial H/dq_i} = - \frac{dq_j}{\partial H/\partial p_j} = dt.$$

Si H ne dépend pas de t, alors H est une intégrale première de (Σ).

On suppose donc que H ne dépend pas de t. Le système différentiel

$$(\Sigma') \qquad \frac{dp_1}{-\partial H/dq_1} = \ldots = \frac{dp_r}{-\partial H/dq_r} = \ldots = \frac{dq_r}{\partial H/\partial p_r}, \quad H = b_0 = C^{te},$$

est défini dans le sous espace V_{n-2} de V_n d'équation $H = b_0$, t = constante. Les résultats précédents permettent de démontrer:

Théorème (d). *La forme $\pi' = \Sigma_i dp_i \otimes dq_i$ induite par π dans V_{n-2} est de classe $r - 1$, π' est un invariant intégral absolu de* (Σ').

On pourra se reporter utilement à [**6**, p. 52-58; **7**, chap. 1, 2, et 4; **5**, p. 121-134].

1.3. Variétés fibrées. Les variétés fibrées considérées dans cet article admettent presque toutes le cercle S_1 comme fibre. On rappelle que V_{n+1} est une variété fibrée de fibre S_1 et de base V_n, si V_{n+1} est muni d'une relation d'équivalence ρ telle que V_n soit l'espace quotient de V_{n+1} par ρ et si l'application canonique P de V_{n+1} sur V_n jouit de la propriété suivante:

Tout point x de V_n admet un voisinage Ω_x dont l'image réciproque $P^{-1}(\Omega_x)$ est homéomorphe au produit topologique $\Omega_x \times S_1$ de Ω_x par S_1, par un homéomorphisme ϕ; de plus si $\bar{x} \in \Omega_x$ et si pr désigne la projection canonique de $\Omega_x \times S_1$ sur Ω_x, la relation suivante est vérifiée: $\mathrm{pr}\, \phi(P^{-1}(\bar{x})) = \bar{x}$.

En fait les espaces fibrés considérés dans la suite admettent une structure encore plus précise: la fibre S est isomorphe au cercle euclidien et le groupe structural de la fibration est le groupe des rotations de S. On a un espace fibré principal. Cependant ces propriétés ne seront pas utilisées dans la suite.

La théorie des variétés fibrées à trois dimensions est exposée dans [**23**]. Ce travail examine aussi le cas des fibres exceptionnelles; ce dernier cas se présente dans certains problèmes de dynamique (cf. §**2.2**). Pour la théorie générale des espaces fibrés on pourra se reporter à [**10**; **11**; **12**]; on trouvera en particulier le lemme du relèvement des homotopies dans [**11**, p. 155] (cf. §**6.4**); dans [**8**] on trouvera une théorie des classes caractéristiques. Il est peut-être utile de rappeler la définition suivante:

Définition. *Une application s de la base V_n dans V_{n+1} est une section de V_{n+1}, si s vérifie la relation $P.s$ = identité.*

II. Quelques exemples simples

2.1. Un exemple classique. L'exemple classique, certainement le plus simple, d'un système différentiel du type envisagé est l'équation du deuxième ordre:

$$(2) \qquad x'' + x = \mu f(x,x')$$

qui pour $\mu = 0$ se réduit à

$$x'' + x = 0. \tag{4}$$

Les lignes intégrales de (4) sont représentées dans le plan des phases (x', x) par les cercles concentriques $x^2 + x'^2 =$ constante. (Ces trajectoires constituent bien une fibration du plan (x, x') auquel on a enlevé l'origine).

La même remarque s'applique au système de type plus général:

$$x'' + g(x) = \mu f(x, x') \tag{5}$$

à condition que les courbes $x'^2 + 2H(x) = C^{te}$ (où $H(x)$ est une primitive de $g(x)$) soient des orbes enveloppant l'origine.

2.2. Petits mouvements. L'étude des petits mouvements au voisinage d'une position d'équilibre stable d'un système dynamique conservatif à q degrés de liberté à liaisons holonomes sans frottement et indépendantes du temps, revient à l'intégration d'un système différentiel du type suivant:

$$dx_i = y_i dt, \qquad dy_i = -\omega_i^2 x_i dt \qquad (i = 1, \ldots, q) \tag{6}$$

où les ω_i sont des constantes réelles non nulles.

Soit R^{2q} l'espace numérique euclidien dans lequel les (x_i, y_i) forment un système de coordonnées cartésiennes; $R^{\circ}{}_{2q}$ désigne l'espace R^{2q} privé de son origine. Si tous les ω_i sont égaux, les trajectoires de (6) forment une fibration de $R^{\circ}{}_{2q}$ dont la base est homéomorphe au produit topologique $R \times P_q(G)$ de la droite numérique R par l'espace projectif complexe $P_q(G)$ à q dimensions complexes [**2**, Livre III, chap. VIII, §**4**) (On remarquera en effet que l'espace numérique euclidien réel R^{2q} peut être considéré comme un espace vectoriel complexe G^q à q dimensions complexes, dans lequel les coordonnées sont: $z_i = x_i + \epsilon y_i$ $(i = 1, \ldots, q;\ \epsilon^2 = -1)$. Les droites D de G^q issues de l'origine, sont représentées dans R^{2q} par des sous-espaces à 2 dimensions. Les traces S des droites D sur la sphère Σ_{2q-1}, de centre O et de rayon 1 dans R^{2q}, sont des grands cercles de Σ_{2q-1}. Les cercles S sont précisément les trajectoires de (6) sur Σ_{2q-1}. On voit donc qu'il y a une correspondance bi-univoque entre les cercles S et les droites D; mais les droites D de G^q sont précisément les points de $P_q(G)$).

Si les ω_i au lieu d'être égaux, sont proportionnels à des nombres rationnels, les trajectoires de (6) sont encore fermées; mais elles ne forment plus une fibration de $R^{\circ}{}_{2q}$. En effet la condition du produit local n'est plus satisfaite Nous avons ici un exemple de structure fibrée avec fibres exceptionnelles [**23**]; en effet il existe un entier k tel que toute trajectoire de (6) possède un voisinage saturé, dont le revêtement à k feuillets est fibré (au sens usuel) par les composantes connexes des images reciproques des trajectoires de (6). On verra facilement, sans qu'il soit nécessaire d'entrer dans les détails, comment la suite pourrait s'appliquer à ce cas un peu plus général.

On peut aussi interpréter le système (6) comme définissant le mouvement de q oscillateurs harmoniques. Si les constantes ω_i sont proportionnelles à des

nombres rationnels, le système différentiel obtenu en ajoutant au deuxième membre de (6) des termes perturbateurs correspond au problème typique de la résonance. (Les périodes fondamentales sont commensurables.)

2.3. Géodésiques d'une sphère S_q. Les géodésiques (ou grands cercles) d'une sphère euclidienne S_q, à q dimensions, sont les projections dans S_q des trajectoires d'un champ de vecteurs E_0 défini dans la variété S^*_{2q-1} des vecteurs unitaires tangents à S_q. Il est évident que les trajectoires de E_0 constituent une fibration de S^*_q dont la variété de base est homéomorphe à la variété des droites (orientées) de l'espace projectif réel P_q à q dimensions. On trouvera d'autres exemples en [**27**].

2.4. Éllipses kepleriennes. Enfin les trajectoires élliptiques d'un point matériel M attiré par un point fixe suivant la loi de Newton, constituent une fibration d'une certaine portion de l'espace des phases R^6. Le problème restreint des trois corps fournit aussi des exemples [**29**].

2.5. Quelques conséquences dues à l'existence d'un invariant intégral. Tous les exemples cités ci-dessus mettent en jeu des systèmes dynamiques admettant l'invariant intégral de M. Elie Cartan [**7**]. On remarque dans tous ces exemples que le temps nécessité pour décrire une trajectoire ne dépend que de la constante des forces vives et non de la trajectoire. Ceci est tout à fait général.

Soit (Σ_0) le système différentiel

$$dy = E_0(y)\,dt, \tag{7} \qquad (\Sigma_0)$$

où E_0 est un champ de vecteurs sur V_{n+1} vérifiant les conditions (c) et (d) de l'Introduction. On suppose que (Σ_0) admet l'invariant intégral relatif $\omega = \pi - H_0dt$, ou π est une forme de Pfaff sur V_{n+1} et H_0 une fonction numérique sur V_{n+1}. Soit $T(y)$ le temps nécessaire pour décrire un tour sur la trajectoire issue de $y \in V_{n+1}$. Le système (Σ_0) est un système différentiel défini sur la variété produit $V_{n+1} \times R$ (où R est l'espace du paramètre t). La transformation Θ de $V_{n+1} \times R$ sur lui-même définie par les relations $y \to y, t \to t + T(y)$ applique chaque trajectoire de Σ_0 sur elle-même. La transformation Θ n'altère pas l'invariant intégral absolu $d\omega$ obtenu en prenant la différentielle extérieure de ω (cf. **1.2**, théorème a'). Or l'image transposée de $d\omega$ par Θ est:

$$\Theta^*(d\omega) = d\pi - dH_0 \otimes dt - dH_0 \otimes dT; \tag{8}$$

d'où on conclut que $dH_0 \otimes dT = 0$, ou encore que T est localement une fonction de H_0. En résumé on peut énoncer:

THÉORÈME I. *Si le champ E_0 vérifie les hypothèses* (c) *et* (d) *énoncées l'Introduction, et si de plus le système différentiel* (**7**) *associé à E_0 admet un invariant intégral relatif de la forme $\omega = \pi(y, dy) - H_0(y)dt$, la période $T(x)$ nécessaire pour décrire un tour sur la trajectoire issue de x est une fonction constante sur les*

composante connexes des variétés de niveau de la fonction $H_0(y)$, [**25**, p. 299; **24**, p. 125].

Si le système différentiel (**7**) admet l'invariant intégral $\omega = \pi - H_0(y)\, dt$, la fonction H_0 est une intégrale première de (**7**). On suppose que V_{n+1} et $H_0(y)$ satisfont à l'hypothèse suivante:

HYPOTHÈSE I. (1) *la forme* dH_0 *est différente de* 0 *en tout point* x *de* V_{n+1} *vérifiant* $H_0(x) = b_0$, *où* b_0 *est une constante donneé.*

(2) *La dimension* $n + 1$ *est paire*; $n + 1 = 2q$ *et* $q > 1$.

(3) *La forme* $d\pi$ *est de classe* $q = \frac{1}{2}(n + 1)$ (cf. **1.2**, définition (b)).

On peut énoncer le théorème suivant:

THÉORÈME II. *Si le système différentiel* (**7**) *admet l'invariant intégral relatif* $\omega = \pi - H_0(y)dt$ *et si de plus on fait l'hypothèse* I, *il existe sur la sous-variété* W_{n-1} *de* V_n *d'équation* $H_0(P^{-1}(x)) = b_0$ *une forme fermée* Ω *de degré deux et de classe maximum.*

En effet soit E'_0 la restriction du champ E_0 à la sous-variété W_n de V_{n+1} d'équation $H_0(y) = b_0$. Le champ E'_0 est un champ de vecteurs sur W_n. Soit π' la forme induite par π dans W_n. La forme $d\pi$ est un invariant intégral absolu du système différentiel associe à E'_0 (cf. **1.2**, théorème (d)):

$$dy \otimes E'_0(y) = 0. \quad (9) \qquad (\Sigma'_0)$$

Par conséquent $d\pi'$ s'exprime uniquement à l'aide d'intégrales premières de (Σ'_0). Or une intégrale première ϕ de (Σ'_0) est localement de la forme $\phi = \tilde{\phi} . P$ où $\tilde{\phi}$ est une fonction sur la base W_{n-1} de W_n. Comme les fibres de W_n sont connexes, on voit que $d\pi'$ est l'image par l'application transposée P^* de P d'une forme extérieure Ω de W_{n-1}: donc $d\pi' = P^*(\Omega)$. Cette dernière relation montre que $d\Omega = 0$ et que Ω a la même classe que $d\pi'$. D'autre part $d\pi'$ est de classe $q - 1$.

On suppose dans la suite que W_n *est compact.* On déduit du théorème II un ensemble de conséquences topologiques pour la variété de trajectoires W_{n-1} et pour l'espace fibré $W_n = P^{-1}(W_{n-1})$. *Les formes fermées* $\Omega, (\Omega)^2, \ldots, (\Omega)^q$, *ne sont pas homologues à zéro.* En effet $(\Omega)^q$ ne s'annule en aucun point de W_{n-1}, son intégrale étendue à W_{n-1} n'est donc pas nulle (cf, **1.1**). Par suite $(\Omega)^q$ n'est pas homologue à zéro. Il en résulte que $\Omega, \ldots, (\Omega)^q$ ne sont pas homologues à zéro. *Les nombres de Betti des dimensions paires de* W_{n-1} *ne sont donc pas nuls.*

La relation $d\pi' = P^*(\Omega)$ *montre que la variété fibrée* W_n *n'admet pas de section* (cf. **1.3**).

Cette dernière affirmation peut être demontrée facilement: soit ϕ une section de W_n; la forme induite par $P^*(\Omega)$ dans $\phi(W_{n-1})$ est identique à $\phi(\Omega)$. Cette forme n'est donc pas homologue à zéro puisque ϕ est un homéomorphisme. Mais d'autre part $P^*(\Omega) = d\pi'$ est homologue à zéro; d'où une contradition.

La relation $d\pi' = P^*(\Omega)$ montre que la classe de cohomologie de Ω est la classe caractéristique de la variété fibrée W_n **[8]**.

L'avant-dernière affirmation peut être énconcée sous une forme bien plus générale: on suppose toujours que le système (Σ_0) admette l'invariant intégral relatif $\omega = \pi - H(y)dt$, et on fait toujours l'hypothèse I; mais on ne suppose plus rien les trajectoires fermées de E_0. Soit W_n la sous-variété de V_{n+1} d'équation $H(x) = b_0$. Dans ces conditions *le champ E'_0 n'admet pas de variété transversale compacte* dans W_n. (Une variété transversale de E'_0 est une variété W_{n-1} à $n - 1$ dimensions, plongée dans W_n, telle qu'en tout point y de W_{n-1} le vecteur $E_0(y)$ n'appartienne pas à l'élément de contact à $n - 1$ dimensions tangent en y à W_{n-1}). L'affirmation résulte de la constatation suivante: on suppose que la variété transversale W_{n-1} existe, la forme $d\pi$ induit dans W_{n-1} une forme extérieure fermée de degré deux et de classe q. Cette forme n'est homologue à zéro; ceci est en contradiction avec le fait que $d\pi$ est homologue à zéro.

Le dernier résultat explique la structure relativement compliquée des surfaces de section avec bords qu'on utilise pour étudier les trajectoires de la dynamique **[4]**.

On remarquera que le théorème II *montre que la variété de base W_{n-1} est une variété symplectique* **[15]**.

III. Propriétés Générales

3.1. Le champ $\bar{E}$ sur V_n. Considérons le système différentiel

$$dy = E_\mu(y)dt \tag{1} \qquad (\Sigma_\mu)$$

satisfaisant aux conditions (a), (b), (c), (d) énoncées dans l'Introduction. Soit E' le champ défini dans V_{n-1} par la relation:

$$E'(y) = \left(\frac{\partial}{\partial\mu} E_\mu(y)\right)_{\mu=0}. \tag{10}$$

Soit $y = y(t)$ la solution de (7) telle que $P(y(t)) = x$. On construit le champ $\bar{E}$ défini dans la variété de base V_n de V_{n+1} par la relation:

$$\bar{E}(x) = \int_{S_x} P(E'(y(t)))dt \tag{11}$$

où $x \in V_n$, $S_x = P^{-1}(x)$.

3.2. L'application ψ_μ; trajectoires simplement fermées. Pour étudier les trajectoires fermées de E_μ il est naturel de suivre la méthode classique **[16]** qui consiste à utiliser une section locale de l'espace fibré V_{n+1}. A cet effet supposons V_n muni d'une subdivision simpliciale. Soit K un sous-complexe fini à n dimensions de V_n, vérifiant la condition (S) suivante:

(S) *Il existe un ouvert O vérifiant $K \subset O \subset V_n$, et un homéomorphisme ϕ deux fois continûment différentiable de rang $n + 1$ de $P^{-1}(O)$ sur $O \times S_1$, tel que les fibres de $P^{-1}(O)$ soient appliquées par ϕ sur les fibres de $O \times S_1$.*

Cette condition (S) est vérifiée en particulier si K est un simplexe, ou plus généralement si K est contractile en un point.

Dans ces conditions on peut introduire dans $P^{-1}(O)$ des coordonnées locales (x,θ) (où $x \in O$ et où $\theta \in S_1$ est en nombre réel défini modulo 2π).

Il existe dès lors $\mu_1 > 0$, tel que pour $\mu \leqslant \mu_1$ les trajectoires de (Σ_μ) issues des points $(x_0 \in K, 0)$ puissent être représentées paramétriquement par les équations:

$$x = \phi_\mu(\theta,x_0); \tag{12}$$

où $\phi_\mu(\theta,x_0) = \phi(\theta,x_0,\mu)$ est une fonction définie pour $0 \leqslant \theta \leqslant 2\pi$ qui vérifie $\phi(0,x_0,\mu) = x_0$.

DÉFINITION. *Une trajectoire simplement fermée de E_μ est une trajectoire qui verifie $\phi(2\pi,x_0,\mu) = x_0$.*

On vérifie que cette définition ne dépend pas des coordonnées particulières (x,θ) choisies dans $P^{-1}(O)$. Elle exprime une propriété géométrique de la trajectoire considérée. Dans la suite on ne considérera que des trajectoires simplement fermées sans s'intéresser aux trajectoires fermées d'une autre nature.

La recherche des trajectoires simplement fermées de E_μ contenues dans $P^{-1}(K)$ revient donc à la recherche des points fixes de la transformation ψ_μ de K dans O définie par:

$$\psi_\mu(x_0) = \phi(2\pi,x_0,\mu). \tag{13}$$

De plus si x_0 est un point fixe isolé de ψ_μ, on peut lui associer son indice [**1**, chap. XIV, §2 et 4; **22**, p. 327] qui est, lui aussi, un invariant.

3.3. Lemme fondamental. La suite consiste en des applications du lemme suivant (qui est classique et simple [**16**; **19**, p. 194-204], et qui met en évidence les relations entre lignes intégrales de E_μ et $\bar{E}$):

LEMME FONDAMENTAL. *On a la relation suivante:*

$$\left(\frac{\partial}{\partial\mu}\,\psi_\mu(x_0)\right)_{\mu=0} = \bar{E}(x_0) \tag{14}$$

Pour démontrer ce lemme, on remarque que le système différentiel (Σ_μ) peut s'écrire (au moyen des coordonnées (x,θ) introduites dans $P^{-1}(O)$), sous la forme suivante:

$$dx = P(E_\mu(x,\theta))dt, \qquad d\theta = \mathfrak{E}_\mu(x,\theta)dt, \tag{15}$$

où $\mathfrak{E}_\mu(x,\theta)$ est la mesure algébrique de la projection du vecteur $E_\mu(x,\theta)$ sur la fibre S_x issue de (x,θ). Le système différentiel (1) peut se mettre sous la forme:

$$dx = P(E_\mu)\,\frac{d\theta}{\mathfrak{E}_\mu}\,.$$

On en déduit par intégration:

(16) $$\phi_\mu(x_0,\bar\theta) = \int_0^{\bar\theta} P(E_\mu(y))\,\frac{d\theta}{\mathfrak{E}_\mu}, \qquad \text{où } y = (\phi_\mu(x_0,\theta);\theta).$$

Bien entendu, pour que l'intégrale du deuxième membre de (16) ait un sens il convient de choisir dans un voisinage $\mathfrak{V}x_0$ de x_0 un système de coordonnées locales (x^i) $(i = 1, \ldots, n)$.

Il est dès lors possible de dériver par rapport à μ:

(17) $$\left(\frac{\partial\psi_\mu}{\partial\mu}\right)_{\mu=0} = \lim_{\mu\to 0}\frac{1}{\mu}\int_0^{2\pi} P(E_\mu(y))\,\frac{d\theta}{\mathfrak{E}_\mu(y)}, \qquad \text{où } y = (\phi_\mu,\theta).$$

$E_\mu(y)$ admet le développement limité (cf. 10):

(18) $$E_\mu(y) = E_0(y) + \mu E'(y) + \ldots,$$

qui entraîne d'ailleurs avec des notations évidentes:

(19) $$\mathfrak{E}_\mu(y) = \mathfrak{E}_0(y) + \mu\mathfrak{E}'(y) + \ldots.$$

En remarquant que $P(E_0(y)) = 0$, on peut donc écrire

(20) $$\left(\frac{\partial\psi_\mu}{\partial\mu}\right)_{\mu=0} = \int_0^{2\pi} P(E'(y_0))\,\frac{d\theta}{\mathfrak{E}_0(y_0)}, \qquad \text{où } y_0 = (x_0,\theta).$$

Le lemme résulte alors immédiatement de (20) et de l'identité suivante entre les paramètres t et θ sur la fibre $P^{-1}(x_0) = S_{x_0}$:

$$d\theta = \mathfrak{E}(y_0)dt.$$

Le lemme montre que ψ_μ admet le développement limité: $\psi_\mu(x) = \psi_0(x) + \mu\bar{E}(x) + \ldots$ où ψ_0 est l'application identique.

3.4. Premières applications du lemme fondamental. On suppose que le complexe K (cf. **3.2**) est un simplexe à n dimensions e_i. Soit ∂e_i le bord de e_i. On suppose que le champ $\bar{E}$ ne s'annule en aucun point de ∂e_i. Dans ces conditions on peut associer au champ $\bar{E}$ un certain indice $I(\bar{E}, e_i)$ [**1**, XIV, §2 et **4**], qui est égal à la somme des indices des points singuliers de $\bar{E}$ dans e_i (au cas où ces points sont isolés). L'indice $I(\bar{E}, e_i)$ est parfaitement déterminé par la restriction de $\bar{E}$ à ∂e_i.

Soit f_μ l'application de V_n dans V_n associée à l'équation:

(21) $$dx = \bar{E}(x)d\mu.$$

(Donc $f_\mu(x_0)$, pour x_0 fixé, est la solution de (21) qui pour $\mu = 0$ prend la valeur x_0.) Soient ψ'_μ et f'_μ les restrictions de ψ_μ et f_μ à ∂e_i. On peut définir l'indice de ψ_μ et de f_μ relativement à e_i [**1**, XIV]; le lemme fondamental admet la conséquence classique:

LEMME. *Il existe $\bar{\mu} > 0$, tel que pour $\mu < \bar{\mu}$ les indices $I(\psi'_\mu)$, $I(f'_\mu)$, $I(\bar{E})$ de ψ_μ, f_μ et $\bar{E}$ relativement à K sont égaux (sous réserve toute fois que $\bar{E}$ ne s'annule en aucun point de ∂e_i).*

Le lemme précédent est une conséquence immédiate du résultat suivant: les champs de vecteurs $\bar{E}$, A_μ/μ et B_μ/μ sont égaux, à des infiniments petits par rapport à μ près; ici A_μ est défini par $A_\mu(x) = \overrightarrow{xx'}$, où $x' = \psi_\mu(x)$ et $x \in \partial e_i$, et B_μ est défini par $B_\mu(x) = \overrightarrow{xx''}$ où $x'' = f_\mu(x)$.

Le lemme fondamental et ce dernier lemme admettent les conséquences suivantes:

THÉORÈME III. *Il existe $\bar{\mu} > 0$ tel que pour $\mu < \bar{\mu}$ on ait les propriétés suivantes:*

(1) *Si toutes les trajectoires simplement fermées de E_μ dans $P^{-1}(e_i)$ sont isolées la somme de leurs indices est $I(\bar{E},e_i)$.*

(2) *Si l'indice $I(\bar{E},e_i)$ de $\bar{E}$ relativement à e_i n'est pas nul, il existe au moins une trajectoire simplement fermée de E_μ dans $P^{-1}(e_i)$.*

(3) *Si $\bar{E}$ ne s'annule en aucun point de e_i le champ E_μ n'a pas de trajectoires simplement fermées dans $P^{-1}(e_i)$.*

(On remarquera que $\bar{\mu}$ dépend de E' et de e_i.)

Le théorème III entraîne à son tour:

THÉORÈME IV. *Si V_{n+1} (et par suite V_n) est compact, il existe $\mu' > 0$ (où μ' dépend de $\bar{E}$) tel que pour $\mu < \mu'$ les trajectoires de E_μ jouissent des propriétés suivantes:*

(a) *Si toutes les trajectoires simplement fermées de E_μ sont isolées, la somme de leurs indices est égale à la caractéristique d'Euler-Poincaré χ de V_n.*

(b) *Si $\chi \neq 0$ le champ E_μ admet des trajectoires simplement fermées. (Sous réserve toute fois que les singularités de $\bar{E}$ soient isolées, ou soient des points internes des simplexes à n dimensions d'une certaine subdivision simpliciale de V_n).*

En effet soient $e_i (i = 1, \ldots, r)$ les simplexes à n dimensions d'une subdivision simpliciale de V_n telle que le bord ∂e_i de e_i ne contienne pas de points singuliers de $\bar{E}$. Le théorème IV résulte du théorème III et de la formule $\Sigma I(\bar{E},e_i) = \chi$. [**1**, XIV, §4, théorème I].

Finalement le théorème III (3) permet d'énoncer:

THÉORÈME V. *Soit Δ l'ensemble (fermé) des points singuliers de $\bar{E}$; on peut supposer que V_n est compact. Pour tout voisinage ouvert U de Δ, il existe $\eta > 0$ tel que pour $\mu < \eta$ les trajectoires simplement fermées de E_μ soient contenues dans $P^{-1}(U)$.*

Soit x_0 un point singulier simple de $\bar{E}$ (c'est à dire que les composantes $\bar{E}_i$ de $\bar{E}$ dans un système de coordonnées locales x_i d'origine x_0 admettent des développements limités

$$\bar{E}_i = \sum a_{ij}x_j + \ldots$$

ou a_{ij} sont des constantes dont le déterminant $D(a_{ij})$ n'est pas nul). Le lemme fondamental permet d'énoncer:

THÉORÈME VI. *Si x_0 est un point singulier simple de $\bar{E}$ il existe $\epsilon > 0$ tel que pour $\mu < \epsilon$, le champ E_μ n'admette qu'une trajectoire simplement fermée contenue dans un voisinage convenable de $P^{-1}(x_0)$.*

L'équation $x = \psi_\mu(x)$ n'admet qu'une solution isolée au voisinage de x_0. En effet cette équation peut encore s'écrire

$$F_\mu(x) = \frac{1}{\mu}\left[(x - x_0) - (\psi_\mu(x) - x_0)\right] = 0$$

et on constate que le Jacobien de F_0 en x_0 n'est pas nul.

REMARQUE. *Si l'ensemble Δ des points singuliers de $\bar{E}$ ne vérifie pas les hypothèses du théorème* IV, *on a cependant par le lemme fondamental des renseignements intéressants sur le comportement des trajectoires de E_μ.*

Il importerait de limiter effectivement les nombres $\bar{\mu}, \eta, \epsilon, \ldots$ pour un champ E_μ donné. Cette question ne sera pas examinée dans ce travail. On trouvera en **5.4** une application des résultats de ce chapitre III.

IV. CAS OÙ LA SYSTÈME (Σ_μ) ADMET UN INVARIANT INTÉGRAL, [28]

4.1. Étude du champ $\bar{E}$ dans ce cas particulier. On suppose que le système différentiel (Σ_μ) défini dans $V_{n+1} \times R$, ou R est l'espace de la variable t, admet un invariant intégral $\omega_\mu = \pi - H_\mu dt$. Ici π designe une forme de Pfaff dont la différentielle $d\pi$ est de classe maximum sur V_{n+1} (π est indépendant de μ); H_μ est une fonction numérique sur V_{n+1} dépendant du paramètre μ. Bien entendu ω_μ est supposé deux fois continûment différentiable *et $n + 1$ est supposé pair* $(n + 1 = 2q)$.

On reprend les notations de (**2.5**) et on fait l'hypothèse I de **2.5** sur ω_0 et V_{n+1}. On rappelle que W_n est la variété d'équation $H_0(y) = b_0$ et que $W_{n-1} = P(W_n)$.

Au moyen d'un choix convenable de coordonnés locales (p_i,q_i) $(i = 1, \ldots, q - 1)$, $p \equiv H_0$ et $\bar{q}$ dans V_{n+1}, l'invariant intégral absolu $d\omega_0$ peut s'écrire sous la forme (cf. **1.2**, théorème (b))

$$d\omega_0 = \textstyle\sum_i dp_i \otimes dq_i + dH_0 \otimes (d\bar{q} - dt) \tag{22}$$

il résulte que les fonctions p_i, q_i, H_0, $\bar{q} - t$, forment un système d'intégrales premières de (Σ_0). Les fonctions p_i, q_i, et H_0 peuvent donc être considérées comme des coordonnées locales dans V_n. La coordonnée $\bar{q}$ définit un paramétrage sur les fibres de V_{n+1}. Soit S_y la fibre de V_{n+1} contenant le point y. Si les coordonnées p_i, q_i, H_0 et $\bar{q}$ sont définies dans un voisinage $\mathfrak{V}_y$ de y, on peut étendre le paramétrage $\bar{q}$ à tout un voisinage saturé $\mathfrak{V}$ de S_y, de telle

façon que $\bar{q} - t$ soit une intégrale première de $\sum_0$. Il convient de remarquer que $\bar{q}$ est défini modulo $T(y)$ (cf. **2.5**, théorème I). Les fonctions p_i, q_i, et H_0 s'étendent d'elles-mêmes à $\mathfrak{B}$ puisqu'elles sont constantes sur les fibres. Par conséquent p_i, q_i, H_0, $\bar{q} - t$ forment un système d'intégrales premières de (Σ_0) dans $\mathfrak{B}$. Il en résulte que la formule (**22**) est valable dans $\mathfrak{B}$.

On peut résumer ceci dans le lemme suivant:

LEMME 1. *Soit $x \in V_n$ et soit $S_x = P^{-1}(x)$ la fibre correspondante. On peut trouver des coordonnés locales p_i, q_i $(i = 1, \ldots, q - 1)$ et H_0 dans un voisinage U_x de x, et une fonction $\bar{q}$ dans $P^{-1}(U_x)$ (définie modulo T) telles que*

$$d\omega_0 = \textstyle\sum_i dp_i \otimes dq_i + dH_0 \otimes (d\bar{q} - dt). \tag{22}$$

CONSÉQUENCE DU LEMME 1. *le système $(\sum_\mu)$ admet l'invariant intégral $d\omega_\mu$ où:*

$$d\omega_\mu = \textstyle\sum_i dp_i \otimes dq_i + dH_0 \otimes d\bar{q} - dH_\mu \otimes dt. \tag{23}$$

Le système Σ_μ s'écrit (dans $P^{-1}(U)$:

$$\begin{aligned} dp_i &= -(\partial H_\mu/\partial q_i)\, dt, & dH_0 &= -(\partial H_\mu/\partial \bar{q})\, dt, \\ dq_i &= (\partial H_\mu/\partial p_i)\, dt, & d\bar{q} &= (\partial H_\mu/\partial H_0)\, dt. \end{aligned} \tag{24}$$

En d'autres termes les composantes du champ E' sont respectivement:

sur l'axe p_i, $-\left(\dfrac{\partial^2 H_\mu}{\partial q_i \partial \mu}\right)_{\mu=0}$; sur l'axe q, $\left(\dfrac{\partial^2 H_\mu}{\partial \mu \partial H_0}\right)_{\mu=0}$;

sur l'axe q_i, $\left(\dfrac{\partial^2 H_\mu}{\partial p_i \partial \mu}\right)_{\mu=0}$; sur l'axe H_0, $-\left(\dfrac{\partial^2 H_\mu}{\partial \mu \partial q}\right)_{\mu=0}$.

En posant $H' = \left(\dfrac{\partial H_\mu}{\partial \mu}\right)_{\mu=0}$ on voit que les composantes du champ $\bar{E}$ sont respectivement:

$$-\int_0^T (\partial H'/\partial q_i) d\bar{q}, \quad \int_0^T (\partial H'/\partial p_i) d\bar{q}, \quad -\int_0^T (\partial H'/\partial \bar{q}) d\bar{q} = 0 \tag{25}$$

où l'intégration est étendue le long des fibres de V_{n+1}.

Or soit r l'une des variables p_i, q_i ou $\bar{q}$; tenant compte du théorème I du paragraphe **2.5** on peut permuter les opérations d'intégration et de dérivation ainsi:

$$\int_0^T (\partial H'/\partial r) d\bar{q} = \frac{\partial}{\partial r} \int_0^T H' d\bar{q}$$

car $\dfrac{\partial T}{\partial r} = 0$. Posons:

$$\bar{H} = \int_0^T H' d\bar{q}. \tag{26}$$

Dans ces conditions $\bar{H}$ est une fonction numérique sur V_n. Dès lors les équations différentielles de $\bar{E}$ sont:

$$(27) \qquad dp_i = -(\partial\bar{H}/\partial q_i)d\bar{q}, \quad dq_i = (\partial\bar{H}/\partial p_i)d\bar{q}, \quad dH_0 = 0.$$

En résumé:

LEMME 2. *Le système* (27) *admet l'invariant intégral relatif* $\bar{\omega} = \sum_i p_i dq_i - \bar{H}dq$ *et l'intégrale première* H_0.

Les lignes intégrales de $\bar{E}$ sont tracées sur les variétés de niveau de la fonction H_0. Les singularités du champ $\bar{E}$ coïncident avec les points critiques de la fonction $\bar{H}$ (cf. (27)). La théorie de Morse [**21**] donne donc des renseignements supplémentaires sur la distribution et la nature des singularités du champ $\bar{E}$.

Cependant (compte tenu de (26)) les points singuliers de $\bar{E}$ dans V_n ne sont pas isolés en général. On ne peut donc pas utiliser tels quels les résultats du chapitre précédent. Mais on peut appliquer le lemme fondamental du chapitre précédent (cf. **3.3**).

Par contre si on se restreint à l'étude des trajectoires qui correspondent à une valeur donnée b_0 de la fonction H_μ les résultats antérieurs seront valables. C'est cette question qui va être examineé dans **4.2**.

4.2. Application des résultats du chapitre III. $H_0(y) = b_0$ est par suite de l'hypothèse sur H_0 en (**4.1**) l'équation d'une variété W_n à n dimensions plongée dans V_{n+1}. Si on interprète H_0 comme fonction numérique sur la base V_n de V_{n+1}, l'équation $H_0(x) = b_0$ définit une variété plongée W_{n-1} à $n - 1$ dimensions dans V_n; la variété W_n est une variété fibrée de base W_{n-1} et de fibre S_1; la projection canonique associée à cette fibration est la restriction de P à W_n. *Pour simplifier l'exposé, on suppose que* W_n *est compact.*

Il existe donc un voisinage U de W_n homéomorphe au produit topologique $W_n \times I$, de W_n par un intervalle ouvert $I =]-1, +1[$ de R^1, dans un homéomorphisme h qui applique le point $y \in W_n$ sur le point $(y,0)$ de $W_n \times I$. On identifiera dans la suite $W_n \times I$ avec U.

La variété $W_{n,\mu}$ d'équation $H_\mu(y) = b_0$ admet, pour les petites valeurs de μ, la représentation paramétrique:

$$(28) \qquad \rho = \Phi(\mu,y) \qquad (\rho \in I)$$

et la fonction $\Phi(\mu,y)$ admet un développement limité par rapport à μ:

$$(29) \qquad \Phi(\mu,\rho) = \mu\Phi_0(y) + \cdots.$$

L'application pr: $(y,\rho) \to y$, est un homéomorphisme de $W_{n,\mu}$ sur W_n. Soit $E_{1,\mu}$ la restriction du champ E_μ (associée à (Σ_μ)) à $W_{n,\mu}$ et soit $E_{2,\mu}$ l'image par pr de $E_{1,\mu}$. Le champ $E_{2,\mu}$ vérifie les hypothèses (a), (b), (c), (d) de l'introduction; on peut donc faire la théorie de III pour ce champ, et définir en particulier un champ $\bar{E}_2$ sur W_{n-1} par une opération analogue à (11).

LEMME 3. *Le champ $\bar{E}_2$ est identique à la restriction du champ $\bar{E}$ à W_{n-1}.*

Soit $y(t)$ une solution de (Σ_0), et soit $x = P(y(t))$.

Le champ $\bar{E}_2$ peut être obtenu par l'opération suivante (cf. (11)):

$$\bar{E}_2(x) = \int_{S_x} P(E'_2(y(t))dt,$$

où

$$E'_2(y) = \left[\frac{d}{d\mu} \operatorname{pr} E_\mu(y, \Phi(\mu, y))\right]_{\mu=0}$$

$$= \operatorname{pr} E'(y) + \left(\frac{\partial}{\partial\rho} \operatorname{pr} E_0(y,\rho)\right)_{\rho=0} \left(\frac{\partial\Phi}{\partial\mu}\right)_{\mu=0}.$$

Or

$$\int_{S_x} P(\operatorname{pr} E'(y(t))dt = \int_{S_x} P(E'(y(t))dt = \bar{E}(x). \tag{30}$$

D'autre part

$$\int_{S_x} \left(\frac{\partial}{\partial\rho} \operatorname{pr} E_0(y,\rho)\right)_{\rho=0} \left(\frac{\partial\Phi}{\partial\mu}\right)_{\mu=0} dt = 0.$$

car dans le cas contraire le champ pr $E_0(y,\rho)$ n'aurait pas de trajectoires fermées au voisinage de $P^{-1}(x)$ pour les petites valeurs du paramètre μ, d'après les résultats de III (cf. théorème III (3)). Le lemme 3 est donc une conséquence immédiate de (30).

4.3. Conséquences du lemme 3. Le lemme 3 montre que pour étudier les trajectoires simplement fermées sur $W_{n,\mu}$ (c'est à dire les trajectoires simplement fermées de $\sum_\mu$ pour lesquelles H_μ prend la valeur b_0) il suffit d'étudier les singularités de la restriction $\bar{E}_1$ du champ $\bar{E}$ à W_{n-1}. On pourrait énoncer des théorèmes tout à fait analogues aux théorèmes III, IV, V, et VI; mais ceci est trop long. Voici simplement quelques remarques à ce sujet:

Les singularités de $\bar{E}_1$ sont les points critiques de la restriction $\bar{H}_1$ de la fonction $\bar{H}$ à W_{n-1} (cf. lemme 2 et (29)). A chaque point critique isolé de $\bar{H}_1$ la théorie de Morse [**21**, p. 85-92] permet d'associer des nombres types (entiers positifs) de chaque dimensions et permet ensuite d'établir des relations entre ces nombres types et la topologie de W_{n-1}. En particulier si W_{n-1} est compact, et si tous les points critiques de $\bar{H}_1$ sont simples [**21**, §8] leur nombre est supérieur à la somme des nombres de Betti de W_{n-1} [**21**, §5]. Il convient de remarquer aussi que puisque W_n est compact la fonction $\bar{H}$ admet nécessairement des points critiques.

On peut donc formuler le théorème suivant:

THÉORÈME VII. *On suppose que le système différential (Σ_μ) admet l'invariant intégral $\omega_\mu = \pi - H_\mu dt$ et que la fonction $\bar{H}_1$, restriction à W_{n-1} de la fonction $\bar{H}$ définie par* (26), *n'admet que des points critiques isolés. Dans ces conditions si $\mu \leqslant \eta$ où $\eta > 0$ est un nombre réel convenable, le système Σ_μ admet des trajectoires simplement fermés tracées dans $W_{n,\mu}$. Le nombre de ces trajectoires*

simplement fermées est au moins égal au nombre des points critiques de $\bar{H}_1$. Comme W_n est compact, la fonction $\bar{H}_1$ admet certainement des points critiques.

La théorie de Morse [**21**] *permet d'étudier la nature de ces points critiques. En particulier si les points critiques de $\bar{H}_1$ sont tous simples, leur nombre est supérieur à la somme des nombres de Betti de W_n.*

Pour achever cette étude il conviendrait d'expliciter les relations entre la nature des points critiques de $\bar{H}_1$ et le comportement des trajectoires simplement fermées associées. A cet effet on peut se reporter à [**4**, chap. III].

V. Applications

5.1. Généralités sur les systèmes conservatifs. Soit (Γ) un système mécanique à liaisons holonomes indépendantes du temps, et désignons par V_q la variété à q dimensions des configurations de ce système. On suppose que les forces connues dérivent d'une fonction de force U_μ indépendante du temps (mais fonction du paramètre μ) et que la force vive $\overline{T}_\mu$ dépend aussi de μ. Pour $\mu = 0$, on suppose que les trajectoires sont périodiques, et que leur période T est une fonction continue, de sorte qu'on puisse appliquer les résultats des chapitres III et IV.

Soit V^*_{2q} la variété à $2q$ dimensions des formes linéaires (ou vecteurs covariants) sur V_q. A tout système de coordonnées locales q_i dans V_q correspond de façon canonique un système de coordonnées locales p_i, q_i dans V^*_{2q}, telles que la forme (p_i,q_i) ait précisément pour expression $\sum p_i dq_i$. En d'autres termes (p_i,q_i) désigne le vecteur covariant attaché au point q_i, ayant pour composantes p_i. L'application Θ de l'espace V'_{2q} des vecteurs (contre-variants) tangents à V_q sur V^*_{2q} définie par $q_i \to q_i$, $q'_i \to \partial\overline{T}_\mu/\partial q'_i$ ne dépend pas des coordonnées particulières q_i. On pose suivant l'usage:

$$H_\mu(\Theta(q_i,q'_i)) = \overline{T}_\mu(q_i,q'_i) - U_\mu(q).$$

Donc $H_\mu(p_i,q_i)$ est une fonction sur V^*_{2q}. Le système différentiel qui définit le mouvement admet l'invariant intégral relatif [**7**] (dans $V^*_{2q} \times R$):

$$\omega_\mu = \sum p_i dq_i - H_\mu dt. \tag{31}$$

Les résultats de IV peuvent être utilisés.

Les mêmes considérations sont valables (avec des précautions évidentes) si les liaisons tout en restant holonomes, dépendent du temps, à condition toutes fois que H_μ ne dépende pas du temps.

5.2. Applications aux géodésiques de la sphère S_q [**28**]. Appliquons ceci au problème des géodésiques de la sphère S_q (cf. **2.3**) munie d'un ds_μ^2 de la forme:

$$ds_\mu^2 = d\sigma^2 + \mu\, d\theta^2$$

où $d\sigma^2$ est la forme quadratique fondamentale de la sphère euclidienne S_q et où $d\theta^2$ est une forme différentielle quadratique quelconque sur S_q. Désignons par $S^*_{2(q-1)}$ la variété des grands cercles orientés de S_q.

La fonction $\bar{H}$ (cf. **4.1**,) prend une même valeur en deux points de $S^*_{2(q-1)}$ correspondants à un même grand cercle de S_q avec ses deux orientations opposées. De la sorte $\bar{H}$ peut être considéré comme une fonction numérique sur la variéte $S^{**}_{2(q-1)}$ des grands cercles, non orientés, de S_q; la variété $S^{**}_{2(q-1)}$ admet $S^*_{2(q-1)}$ comme revêtement à deux feuillets.

Les nombres de Betti modulo 2 de $S^{**}_{2(q-1)}$ sont connus [**13**] et peuvent être consignés dans le tableau suivant, dont la loi de formation est évidente: (b_p désigne le nombre de Betti pour la dimension p).

	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8
$q = 2$	1	1	1						
$q = 3$	1	1	2	1	1				
$q = 4$	1	1	2	2	2	1	1		
$q = 5$	1	1	2	2	3	2	2	1	1
...	.	.	.	.	.	.	.	.	.

La somme des nombres de Betti modulo 2 de $S^{**}_{2(q-1)}$ est égale à

$$\tfrac{1}{2}\, q(q+1).$$

Si les singularités de $\bar{H}$ sont simples, alors les géodésiques orientées simplement fermées pour les petites valeurs de μ sont [**21**] *au nombre de* $q(q+1)$ au moins.

Le calcul explicite de la fonction $\bar{H}$ est assez simple ici, aussi allons nous l'indiquer en détail. Remarquons que la fonction de Hamilton H_μ associée à notre problème de géodésiques est définie sur la variété S^*_{2q} des formes linéaires sur S_q. La valeur de H_μ au point z de S^*_{2q} est égale au carré scalaire de la forme z.

Pour avoir des paramètres commodes, nous supposons que S_q est définie dans l'espace euclidien R^{q+1} par l'équation cartésienne:

$$x_1^2 + x_2^2 + \ldots + x_{q+1}^2 = 1.$$

Les formes sur S_q peuvent être représentées par les vecteurs tangents à S_q, dont les composantes seront désignées par $u_i (i = 1, \ldots, q+1)$. Au vecteur (x_i, u_i) correspond la forme induite dans S_q par la forme $\Sigma u_i dx_i$. L'élément d'arc ds_μ^2 est l'élément d'arc induit dans S_q par la forme quadratique suivante définie dans R^{q+1}:

$$ds_\mu^2 = (dx_1^2 + \ldots + dx_{q+1}^2) + \mu \sum \epsilon_{ij} dx_i dx_j$$

où les ϵ_{ij} sont les composantes d'un tenseur deux fois covariant défini dans R^{q+1}. On vérifie que la fonction H_μ admet le développement limité (suivant μ):

$$H_\mu = (u_1^2 + \ldots + u_{q+1}^2) - \mu \sum \epsilon_{ij} u_i u_j + \ldots .$$

Les équations paramétriques des trajectoires de H_0 sont:

(32) $$x_i = A_i \cos t + B_i \sin t, \quad u_i = -A_i \sin t + B_i \cos t;$$

où A_i et B_i sont des constantes qu'on peut interpréter comme coordonnées

d'un point de S^*_{2q}. L'expression de $\bar{H}$ (ou plus exactement de $\bar{H}\,.\,P$) est donnée par:

$$\bar{H}(A_i,B_i) = -\int_0^{2\pi} \sum \epsilon_{ij} u_i u_j dt \tag{33}$$

où on a substitué à u_i et x_i les expressions (32).

En particulier si les ϵ_{ij} sont des polynomes en (x_i,u_i) on voit que $\bar{H}$ est un polynome en (A_i,B_i) ne contenant que des termes de degré pair. Si nous nous limitons aux trajectoires vérifiant $H_0 = 1$, les coordonnées (A_i,B_i) vérifient les relations:

$$\sum A_i^2 = 1, \quad \sum B_i^2 = 1, \quad \sum A_iB_i = 0.$$

5.3. Application au cas particulier d'un éllipsoïde. Un cas particulier remarquable est celui où les coefficients ϵ_{ij} sont des constantes, cas qui correspond à l'éllipsoide à q dimensions. On peut toujours faire en sorte que $\epsilon_{ij} = 0$ si $i \neq j$. Soit $\alpha_i = \epsilon_{ij}$. Si les α_i sont tous distincts, on a le cas d'un éllipsoide d'axes inégaux.

L'integration (33) donne:

$$\bar{H}\,.\,P = -\pi\sum \alpha_i(A_i^2 + B_i^2).$$

Les points critiques de $\bar{H}\,.\,P$ (dans la variété $H_0 = 1$) sont les points où la matrice suivante est de rang inférieur à 4:

$$\begin{pmatrix} \alpha_1 A_1 \ldots \alpha_{q+1} A_{q+1} & \alpha_1 B_1 \ldots \alpha_{q+1} B_{q+1} \\ 0 \quad \ldots \quad 0 & B_1 \ldots \quad B_{q+1} \\ B_1 \quad \ldots \quad B_{q+1} & A_1 \ldots \quad A_{q+1} \\ A_1 \quad \ldots \quad A_{q+1} & 0 \ldots \quad 0 \end{pmatrix}.$$

En effet cette dernière condition exprime la nullité de la forme:

$$d\bar{H} \otimes d(\sum A_i^2) \otimes d(\sum B_i^2) \otimes d(\sum A_iB_i) = 0.$$

Si les α_i sont tous distincts, on met en évidence les $\frac{1}{2}q(q+1)$ groupes de points critiques obtenus en annulant toutes les coordonnées $B_1, \ldots, B_{q+1}$; $A_1, \ldots, A_{q+1}$ sauf une de chaque groupe, les deux coordonnées non nulles devant avoir des indices différents. Par exemple on peut faire:

$$B_1 = 0, \ldots, B_q = 0; \qquad A_2 = 0, \ldots, A_{q+1} = 0.$$

On vérifie qu'on obtient ainsi tous les points critiques de $\bar{H}$. Ainsi on a mis en évidence, pour les petites valeurs de μ, les éllipses principales de l'éllipsoide à axes inégaux comme seules géodésiques simplement fermées.

5.4. Application aux oscillations de relaxation. Soit (Σ_μ) le système différentiel (cf. (6)):

$$dx_i = u_i dt, \quad du_i = (-x_i + \mu f_i(x,u))dt; \tag{34}$$

défini dans R^{2q}. Les champs E_μ, E_0 et E' ont respectivement pour composantes: $(u_i, -x_i + \mu f_i(x,u))$, $(u_i, -x_i)$, $(0, f_i(x,u))$.

Les lignes intégrales de E_0 ont pour équation:

$$x_i = A_i \cos t + B_i \sin t \qquad u_i = -A_i \sin t + B_i \cos t, \tag{35}$$

où A_i et B_i sont des constantes.

Soit $\mathfrak{F}(x,u) = \Sigma_i(x_i^2 + u_i^2)$ «l'énergie» du système (Σ_μ). La fonction $\mathfrak{F}(x,u)$ est constante sur les trajectoires de E_0; elle est donc une fonction définie l'espace des fibres de E_0. On vérifie facilement que:

$$(d\mathfrak{F}, \bar{E}) = 2\int_0^{2\pi} \Sigma_i u_i f_i(x,u)\, dt,$$

où on a substitué à x_i et u_i les expressions (35). (On désigne par (ω, v) le produit scalaire de la forme ω par le vecteur v.)

Soit Δ_{2q} les sous-espaces de R^{2q} défini par les inégalités:

$$m \leqslant \mathfrak{F}(x,u) \leqslant M. \tag{36}$$

L'espace Δ_{2q} est fibré par les trajectoires de E_0 et l'espace de base de Δ_{2q} est isomorphe à $P_q(G) \times I$ (où I est l'intervalle fermé $[\sqrt{m}, \sqrt{M}]$) (cf. **2.2**).

Dans de nombreux problèmes de dynamique on a la circonstance suivante (pour des valeurs convenables de M et m):

$$(d\mathfrak{F}, \bar{E}) < 0 \text{ sur } P_q(G) \times \{\sqrt{M}\} \text{ et } (d\mathfrak{T}, \bar{E}) > 0 \text{ sur } P_q(G) \times \{\sqrt{m}\}, \tag{37}$$

dont l'interprètation dynamique est évidente (cf. **6.4**).

Si la relation (37) a lieu, le champ $\bar{E}$ sur $P_q(G) \times I$ ne s'annule en aucun point du bord de $P_q(G) \times I$ et il est dirigé vers l'intérieur de $P_q(G) \times I$. On en conclut que $\bar{E}$ a des singularités, dont la somme des indices est $\chi(P_q(G)) = q$. Dans le cas particulier où $q = 1$, on retrouve les résultats classiques sur l'équation (2) [**19**, p. 194-204].

VI. Sur l'existence de solutions périodiques sous des conditions plus larges

6.1. Introduction. Les résultats exposés dans les chapitres précédents présentent un *grave inconvénient*: le paramètre μ est non seulement petit, mais encore il est borné par $\epsilon > 0$ où ϵ dépend de E'. On se propose de voir comment on peut démontrer l'existence de solutions périodiques sous des conditions un peu plus larges.

H. Seifert a démontré un théorème dans ce sens [26]:

Théorème. *Soit E_0 un champ de vecteurs sans singularités, défini sur une variété de Riemann compacte V_3 à trois dimensions, dont les trajectoires sont fermées et forment une fibration de V_3 dont la base est V_2. Soit E un deuxième champ de vecteurs sur V_3 vérifiant en tout point $y \in V_3$:*

$$||E_0(y) - E(y)|| \leqslant \epsilon$$

(où $\epsilon > 0$ est un nombre réel convenable attaché au couple (E_0, V_3) et où $||E(x)||$ est la norme de $E(x)$). *Dans ces conditions, si la caractéristique d'Euler-*

Poincaré χ de V_2 n'est pas nulle, le champ E admet des trajectoires simplement fermées; de plus la somme de leurs indices (si ces trajectoires sont en nombre fini) *est égale à χ.*

Ce théorème admet des applications à l'étude des géodésiques sur S_2 (voir aussi **6.3**). Un théorème du même type (mais bien plus facile) sera démontré dans la suite, et le résultat sera appliqué aux oscillations de relaxation.

6.2. Un résultat général. Soit V_n une variété compacte à n dimensions, sur laquelle est défini un champ de vecteurs E_0, sans singularités, dont les trajectoires sont fermées et forment une fibration de V_n. Pour plus de commodité on suppose que V_n est doué d'une structure de variété de Riemann et on désigne par $\|E(x)\|$ la norme du vecteur $E(x)$.

A tout $x \in V_n$ on peut associer la plaque $Q(x)$ engendrée par les arcs de géodésiques issues de x, normaux à $E(x)$ et de longueur inférieure à un nombre positif $\bar{a}$ donné. Si $\bar{a}$ est choisi convenablement $Q(x)$ est homéomorphe, par une application h, à une boule compacte B_{n-1} à $n-1$ dimensions de R^{n-1} et les géodésiques issues de x sont appliquées par h sur les rayons de B_{n-1}.

Soit E un deuxième champ de vecteurs défini sur V_n et vérifiant $\|E(x) - E_0(x)\| < \epsilon$ pour tout $x \in V_n$ (ici ϵ est un nombre positif donné). On peut choisir $\epsilon > 0$ de sorte que la trajectoire de E issue de x recoupe $Q(x)$ (pour la première fois) en x', et de sorte que le vecteur $E'(x)$ tangent en x à la géodésique orientée xx', issue de x et allant vers x', ayant pour longueur de xx', soit une fonction continue de x.

L'existence d'une trajectoire simplement fermée du champ E, est équivalente à l'existence de singularités du champ E' qui vient d'être défini. En particulier, si la variété V_n n'admet pas deux champs de vecteurs linéairement indépendants en tout point de V_n, le champ E admet nécessairement une trajectoire simplement fermée [**20**, d].

En particulier on sait que la sphère S_{4s+1} à 4^{s+1} dimensions (s entier) n'admet pas deux champs de vecteurs linéairement indépendants [**10**].

6.3. Une application de 6.2 à un système dynamique. Soit (Γ) un système dynamique, à liaisons holonomes et indépendantes du temps, dont l'espace de configuration est l'espace numérique R^n des variables $q_i (i = 1, \ldots, n)$ et dont l'espace de phase est l'espace numérique R^{2n} des variables, q_i et q'_i. On fait sur (Γ) les hypothèses suivantes:

(α) *La force vive $2T$ de* (Γ) *a pour expression*:

$$2T = \sum_i (q'_i)^2 + \sum a_{ij}(q')q'_i q'_j \quad \text{(on pose } \sum (q'_i)^2 = 2T_0\text{)}.$$

(β) *Les forces données appliquées à* (Γ) *se répartissent dans les trois catégories suivantes*:

(1) *Des forces dérivant de la fonction de force $U(q) = \sum -(\omega_i q_i)^2$ où les ω_i sont des constantes proportionnelles à des nombres rationnels* (*cf.* **2.2**).

(2) *Des forces perturbatrices dérivant de la fonction de force* $B(q)$.

(3) *Des forces perturbatrices (de nature électro-magnétique par exemple) dont le travail élémentaire réel est nul, mais dont le travail virtuel dans le déplacement élémentaire* δq *à partir de l'état* (q_i, q'_i) *a pour expression*:

$$\delta W = \sum b_i(q_i, q'_i)\, \delta q_i.$$

Dans ces conditions on vérifie que le système (Γ) admet l'intégrale première des forces vives: $(T - U - B) = h = \text{Cte}$. En particulier si les termes perturbateurs B, a_{ij}, b_i sont nuls, l'intégrale première des forces vives s'écrit: $T_0 - U = b$. Cette équation définit dans R^n un sous-espace homéomorphe à la sphère S_{2n-1}; de plus les trajectoires de (Γ) tracées sur S_{2n-1} forment une fibration de S_{2n-1} (cf. **2.2**).

On déduit facilement de **6.2**, sans qu'il soit nécessaire d'entrer dans détails:

THÉORÈME. *On suppose que les fonctions* $|a_{ij}|$, $|B(q)|$ *et* $|b_i|$ *sont bornées par* $\epsilon > 0$ *sur le sous espace de* R^{2n} *défini par l'inégalité* $\frac{1}{2}b \leqslant T_0 - U \leqslant 2b$. *Il existe un nombre réel* $\eta > 0$, *ne dépendant que de* b *et* ω_i, *tel que si* $\epsilon < \eta$, *et si* n *est impair le système* (Γ) *admette au moins une trajectoire simplement fermée pour laquelle la constante des forces vives est* h.

6.4. Applications aux oscillations de relaxation. Envisageons de nouveau le système différentiel: (voir (6))

$$dx_i = u_i dt, \qquad du_i = (-x_i + f_i(x,u))dt, \qquad (i = 1, \ldots, q). \tag{38}$$

Ce système a une signification mécanique ou physique évidente. Les fonctions f_i sont des fonctions perturbatrices. Les trajectoires du système (38) non perturbé (c'est à dire où l'on a fait $f_i = 0$) forment une fibration de l'espace R^{2q} pointé en 0; de plus le système admet l'intégrale première:

$$F(x_i, u_i) = \sum_i (u_i^2 + x_i^2).$$

De plus on admet, ce qui est naturel dans l'interprétation dynamique, que les fonctions perturbatrices sont telles que l'hypothèse suivante $\mathfrak{H}$ soit vérifiée:

HYPOTHÉSE $\mathfrak{H}$: *Soit* E_0 *le champ associé au système* (6). *Soit* E' *le champ défini en* **6.2**, *normal à* E_0, *sur la sphère creuse* Δ_{2q} *d'équation*:

$$m \leqslant F \leqslant M$$

(où $M > m > 0$ *sont deux nombres réels convenables*). *Soit* E'' *la restriction du champ* E' *au bord de la sphère creuse* Δ_{2q}. *Le champ* E'' *est dirigé vers l'intérieur de* Δ_{2q}.

(Bien entendu pour pouvoir construire le champ E' on suppose que les fonctions perturbatrices f_i sont assez petites; c'est à dire qu'elles vérifient des inégalités du type:

$$|f_i(x,u)| < \epsilon_i$$

où les constantes positives ϵ_i ne dépendent que de m et M). L'hypothèse $\mathfrak{H}$ revient en gros à exiger que l'énergie F diminue aux grandes vitesses et augmente aux petites vitesses (cf. **5.4**). En tout cas il est facile en général de vérifier si cette hypothèse est satisfaite.

THÉORÈME VIII. *Si le système* (38) *vérifie* $\mathfrak{H}$, *et si de plus q est impair, le système* (38) *admet au moins une solution périodique.*

Pour démontrer VIII on démontre que si $2q - 1 = 4^{s+1}(s \geqslant 2)$ le champ E' a des singularités. On peut d'ailleurs supposer que la restriction du champ E' au bord de Δ_{2q} est normale au bord de Δ_{2q} et que $||E'(x)|| = 1$ en tout point x de Δ_{2q} qui n'est pas un point singulier de E'. Le théorème VIII est donc équivalent au théorème géométrique suivant:

Soit S_{4s+1} la sphère de rayon 1 *de l'espace euclidien R^{4s+2} (s entier, $s > 1$). Dans ces conditions le champ des normales extérieures de S_{4s+1} ne peut pas être déformé dans le champ des normales intérieures, si pendant la déformation les vecteurs du champ déformé sont assujettis à rester orthogonaux aux vecteurs d'un champ E'_0 donné tangent à S_{4s+1}.*

Dans le produit topologique $S_n \times I$ (où I est l'intervalle fermé $[-1 + 1]$) on considère donc un champ de vecteurs E_0 qui en tout point $(x_0,\rho_0) \in S_n \times I$ est tangent à la sphère $\rho = \rho_0$ et qui vérifie $E_0(x,\rho) = E_0(x,\rho')$ quels que soient ρ et ρ'. Soit E' un deuxième champ de vecteurs sur $S_n \times I$, qui sur le bord $\overline{S}_n$ de $S_n \times I$ est orthogonal à $\overline{S}_n$, et qui est dirigé vers l'intérieur de $S_n \times I$.

Il s'agit de reconnaître s'il existe un prolongement E'' du champ E', déjà défini sur le bord de $S_n \times I$, à $S_n \times I$ tel que la condition d'orthogonalité entre E'' et E_0 soit satisfaite.

Soient respectivement H et K les deux hémisphères compacts de S_n relativement à un plan équateur donné. Les deux lemmes suivants (1) et (2) sont des conséquences immédiates du théorème du relèvement des homotopies [**11**]:

(1) *On peut prolonger E' au sous-espace $K \times I$ de sorte que la relation d'orthogonalité avec E_0 soit satisfaite.*

(2) *Si le prolongement désiré de E' à $S_n \times I$ existe, il est possible de prolonger tout champ $\overline{E}'$ sur $K \times I$ en un champ E'' sur $S_n \times I$ vérifiant la condition d'orthogonalité (à condition que $\overline{E}'$ soit un prolongement de E' à $K \times I$ et vérifie la condition d'orthogonalité).*

On peut supposer que le prolongement $\overline{E}'$ de E' à $K \times I$ vérifie:

$$\overline{E}'_1(x,\rho) = (1 - |\rho|)E'_1(x, 0), \quad \overline{\mathfrak{E}}'(x,\rho) = -\rho \tag{39}$$

où $x \in S_n$, $\rho \in I$, où $\overline{E}'_1$ est la composante de $\overline{E}'$ selon S_n, et $\overline{\mathfrak{E}}'$ la composante suivant I. Dans ces conditions la restriction de $\overline{E}'$ à $\dot{H} \times I$ (où $\dot{H}$ est le bord de H) satisfait aussi aux relations (39) (où cette fois-ci $x \in \dot{H}$).

Soit Q l'espace des vecteurs non nuls de $H \times I$, orthogonaux au champ E_0. On a évidemment un parallélisme entre les vecteurs de Q attachés à des points

se projetant sur un même point de H. D'autre part on peut définir un parallélisme parmi les vecteurs de Q tangents à H, (en effet H est contractile en un point). Il en résulte un parallélisme naturel pour tous les vecteurs de Q.

A tout vecteur de Q on peut associer grâce à parallélisme un point de la sphère S_{n-1} à $n-1$ dimensions; appelons γ l'application continue ainsi définie de Q dans S_{n-1}.

Si le prolongement désiré du champ E' à $S_n \times I$ existe, l'application θ du bord $\partial(H \times I)$ de $H \times I$ dans S_{n-1} définie en associant à $x \in \partial(H \times I)$ le vecteur du champ E' (ou $\bar{E}'$) et en prenant son image par γ, est homotope à zéro. Or la restriction de θ à $\dot{H} \times \{0\}$ est une application θ' de $\dot{H} \times \{0\}$ dans l'équateur S_{n-2} de S_{n-1}. D'après (39), θ n'est autre chose que la Einhängung de Freudenthal [**14**] de θ', et est donc (dans notre cas particulier) essentielle si et seulement si θ' est essentielle. Or on sait d'après la théorie classique que si $n = 4q + 1$, l'application θ' est essentielle (donc aussi θ); d'où le théorème. [**14**; **10**].

Dans le cas particulier où $q = 1$ le théorème reste valable et redonne des résultats bien connus dans le cas des oscillations de relaxation à un degré de liberté. [**19**, p. 184-194].

Références

1. P. Alexandroff et H. Hopf, *Topologie* (Berlin, 1935).

2. N. Bourbaki, *Éléments de mathématique* (Paris, 1940, 1942, 1947, 1948).

3. P. Bidal et G. de Rham, *Les formes différentielles harmoniques*, Commentarii Mathematici Helvetici, vol. 19 (1947), 1-49.

4. G. D. Birkhoff, *Dynamical systems*, Amer. Math. Soc. Colloq. Publ., vol. 9 (1927).

5. C. E. Carathéodory, *Variationsrechnung und partielle Differential-gleichungen* 1 *Ordnung* (Leipzig, 1935).

6. É. Cartan, *Les systèmes différentiels extérieurs* (Paris, 1946).

7. ———, *Leçons sur les invariants intégraux* (Paris, 1922).

8. S. Chern, *Characteristic Classes of Hermitian Manifolds*, Ann. of Math., vol. 47 (1946), 85-121.

9. C. Chevalley, *Theory of Lie Groups I* (Princeton, 1946).

10. B. Eckmann, *Systeme von Richtungsfeldern in Sphären und stetige Lösungen komplexer linearer Gleichungen*, Commentarii Mathematici Helvetici, vol. 15 (1942), 1-26.

11. ———, *Zur Homotopietheorie gefaserter Räume*, Commentarii Mathematici Helvetici, vol. 14 (1941), 141-192.

12. Ch. Ehresmann, *Sue la théorie des espaces fibrés*, Colloque de topologie algébrique (Paris, 1948).

13. ———, *Sur la topologie de certaines variétés algébriques*, Jour. de math., 104 (1937), 69-100.

14. H. Freudenthal, *Über die klassen der Sphärenabbildungen*, Compositio Mathematica, vol. 5 (1934), 299-314.

15. H. Guggenheimer, *Sur les variétés qui possèdent une forme quadratique fermée*, Comptes Rendus, vol. 232 (1951), 490.

16. J. Haag, *Sur la synchronisation des systèmes à plusieurs degrés de liberté*, Ann. Sci. de l'École Normale Sup., vol. 64, (1947) 237-338.

17. ———, *Sur la synchronisation des systèmes oscillants non linéaires*, Ann. Sci. de l'École Normale vol. Sup., 67 (1950), 321-392.

18. N. Kryloff et N. Bogoliouboff, *An Introduction to Non-Linear Mechanics* (Princeton, 1948).

19. S. Lefschetz, *Lectures on differential equations* (Princeton, 1946).

20a. G. Reeb, *Variétés de Reimann dont toutes les géodésiques sont fermées*, Bull. de la classe des Sci., Bruxelles, vol. 36 (1950), 324-329.

20b. ——— Comptes Rendus, Paris, vol. 229 (1949), 969-971.

20c. ——— Comptes Rendus, Paris, vol. 228 (1949), 1097-1098 et 1196-1198.

20d. ——— Archiv der Mathematik, vol. 50 (1949), 205-206.

21. H. Seifert et W. Threlfall, *Variationsrechnung im Grossen*, (Leipzig, 1938).

22. ———, *Lehrbuch der Topologie*, (Leipzig, 1934).

23. H. Seifert, *Topologie dreidimensionaler gefaserter Räume*, Acta Math., vol. 60 (1932), 147-238.

24. O. Zoll, *Über geschlossene geodätische Linien*, Math. Ann., vol. 57 (1903), 108-133.

25. P. Funk, *Über Flächen mit lauter geschlossenen geodätischen Linien*, Math. Ann., vol. 74 (1913), 278-300.

26. H. Seifert, *Closed integral curves in 3-space, ...*, Proceedings Am. Math. Soc., vol. 1 (1950), 287-302.

27. É. Cartan, *Sur certaines formes riemanniennes ...*, Ann. Sci. de l'École Normale Sup., vol. 44 (1927), 466-467.

28. H. Poincaré, *Sur les lignes géodésiques des surfaces convexes*, Trans. Am. Math. Soc., vol. 6 (1905), 237-274.

29. A. Speiser, *Topologische Fragen aus der Himmelsmechanik*, Vierteljahresschrift der Naturforsch. Gesellsch. i. Zürich, vol. 85 (1940), 204-213.

Saverne, France

NORMAL SAMPLES WITH LINEAR CONSTRAINTS AND GIVEN VARIANCES

D. A. S. FRASER

1. Summary. In Biometrika (1948) a paper [1] by H. L. Seal contained a theorem applying to "n random variables normally distributed about zero mean with unit variance, these variables being connected by means of k linear relations."[1] Arising from this is the question of how to obtain a set of normal variates connected by k linear relations and such that each variate has unit variance: or, more generally, connected by k linear relations and such that each variate has a given variance. The procedure for obtaining such a set of variates when existent from a set of independent normal deviates with unit variances is given in §5. In §§2, 3 and 4, we shall consider various conditions necessary for the existence and construction of such a set.

2. Normal distribution in a linear subspace. Let $(x_1, x_2, \dots, x_n)$ designate a point in an n-dimensional Euclidean space R^n. A set of variates $x_1, x_2, \dots, x_n$ can be considered as a random point in R^n. In the present problem we shall assume the set has a multivariate normal distribution.

Consider k homogeneous and independent linear relations

$$\sum_{j=1}^{n} a_{pj}x_j = 0 \qquad (p = n - k + 1, \dots, n).$$

A variate satisfying these relations and these only will belong to an $(n - k)$-dimensional linear subspace. By taking linear combinations of the above k relations, an equivalent set of k relations can be obtained such that they are orthogonal and normalized:

$$\sum_{j=1}^{n} b_{pj}x_j = 0 \qquad (p = n - k + 1, \dots, n),$$

and

$$\sum_{j=1}^{n} b_{pj}b_{qj} = \delta_{pq}.$$

By adding $n - k$ rows, the matrix $\|b_{pj}\|$ can be completed to an n by n matrix $\|b_{ij}\|$ $(i, j, = 1, 2, \dots, n)$ satisfying the orthogonality conditions

$$\sum_{k=1}^{n} b_{ik}b_{jk} = \delta_{ij}.$$

Received July 8, 1950.

[1]It is to be noted that the statement of the theorem in [1] is incorrect. The theorem applies to the *residuals* of n normal variables after fitting k linear constraints.

This matrix can now be considered as the matrix of an orthogonal rotation of n-space. Consider coordinates $y_1, y_2, \ldots, y_n$ with respect to the new axes: then

$$y_i = \sum_{j=1}^{n} b_{ij}x_j.$$

Since normality is invariant under linear transformations, a set of normally distributed x variates yields a set of normally distributed y variates, and conversely. A set of variates $(x_1, x_2, \ldots, x_n)$ satisfying k linear relations

$$\sum_{j=1}^{n} a_{pj}x_j = 0$$

is transformed by the above to set a of variates $(y_1, y_2, \ldots, y_{n-k})$ satisfying no linear constraints where $y_{n-k+1}, \ldots, y_n$ are identically zero.

3. Conditions on the variance. We have solved the problem of linear constraints by working in an $(n - k)$-dimensional subspace. How do we interpret in this subspace the original variance conditions:

$$\operatorname{var}\{x_i\} = v_i \qquad (i = 1, 2, \ldots, n),$$

with

$$y_r = \sum_{j=1}^{n} b_{rj}x_j,$$

$$x_i = \sum_{r=1}^{n-k} b_{ri}y_r.$$

Each x_i is seen to be a linear combination of the $n - k$ variates y_r and consequently the variance of x_i can be expressed in terms of the elements of the variance covariance matrix of the y_r. Consider now a multivariate normal distribution in the subspace with covariance matrix $\|\tau_{rs}\|$ with respect to the axes $y_1, y_2, \ldots, y_{n-k}$.

Thus the variance conditions after rotation into the subspace become

$$\sum_{r,s=1}^{n-k} b_{ri}\tau_{rs}b_{si} = v_i \qquad (i = 1, 2, \ldots, n).$$

4. Existence. Our problem has now reduced itself to that of finding a multivariate normal distribution in $n - k$ dimensions with covariance matrix $\|\tau_{rs}\|$ such that

$$\sum_{r,s=1}^{n-k} b_{ri}\tau_{rs}b_{si} = v_i \qquad (i = 1, 2, \ldots, n),$$

or

$$\sum_{r,s=1}^{n-k} c_{ri}\tau_{rs}c_{si} = 1 \qquad (i = 1, 2, \ldots, n)$$

where $c_{ri} = v_i^{-\frac{1}{2}}b_{ri}$.

We have n equations with $\binom{n-k+1}{2}$ unknowns. If $\binom{n-k+1}{2} \geqslant n$, a solution will exist and can best be obtained by solving the equation directly. If $\binom{n-k+1}{2} < n$, an application of linear regression theory would be indicated.

To find a matrix $||\tau_{rs}||$, if one exists, is equivalent to finding a generalized ellipsoid

$$\sum_{r,s=1}^{n-k} z_r \tau_{rs} z_s = 1$$

passing through the n points

$$(c_{1i}, \dots, c_{n-k,i}) \qquad (i = 1, 2, \dots, n).$$

This is accomplished using linear regression theory by fitting to the constant 1 the functions $z_r z_s$ $(r,s = 1, 2, \dots, n-k)$ for the n "sample" values given above of the vector $(z_1, z_2, \dots, z_{n-k})$. If the sum of squares for residuals is zero then a quadratic surface exists. However, to have a solution to our distribution problem, the matrix of the quadratic form must be positive. If it is not positive definite, then our variance conditions have imposed a further linear constraint on the set of variates.

5. Conclusions. The problem may be stated: to find normal variables $x_1, x_2, \dots, x_n$ satisfying k homogeneous and independent linear relations

$$\sum_{j=1}^{n} a_{pj} x_j = 0 \qquad (p = n - k + 1, \dots, n),$$

and with

$$\text{var}\{x_i\} = v_i \qquad (i = 1, 2, \dots, n).$$

The solution can be described in five steps.

5.1. Find a matrix $||b_{pj}||$ with $p = n - k + 1, \dots, n$ and $j = 1, 2, \dots n$ with orthogonal and normalized rows equivalent to $||a_{pj}||$ as described in §2.

5.2. Complete $||b_{pj}||$ to an orthogonal matrix $||b_{ij}||$ $(i, j = 1, 2, \dots, n)$.

5.3. Find a quadratic equation

$$\sum_{r,s=1}^{n-k} z_r \tau_{rs} z_s = 1$$

satisfied by the n points

$$(b_{1i} v_i^{-\frac{1}{2}}, \dots, b_{n-k,i} v_i^{-\frac{1}{2}}) \qquad (i = 1, 2, \dots, n),$$

if it exists, directly or by regression theory as in §4. If the equation does not exist or if it exists with a non-positive matrix then the problem has no solution.

5.4. If the matrix $||\tau_{rs}||$ is positive, then find random variates $y_1, y_2, \dots, y_{n-k}$ with zero means and $||\tau_{rs}||$ as covariance matrix. (If $||\tau_{rs}||$ is positive

definite, take the square root matrix of $\|\tau_{rs}\|$ and apply as a linear transformation to $n - k$ independent normal variates with means 0 and variances 1. If positive but not definite, then the previous method will work in a subspace of $y_1, y_2, \ldots, y_{n-k}$.)

5.5. Obtain the set of x variates, thus solving the problem, by applying the transformation

$$x_i = \sum_{r=1}^{n-k} b_{ri} y_r$$

to the y variates obtained in 5.4.

REFERENCES

1. H. L. Seal, *A note on the χ^2 smooth test*, Biometrika, vol. 35 (1948), 202.

University of Toronto

ON A THEOREM OF AUBRY-THUE

ALFRED BRAUER AND R. L. REYNOLDS

1. Introduction. In 1913 L. Aubry [**1**] proved the following theorem:

If a and m are relatively prime, $m > 0$, and if $b/m^{1/2}$ is not an integer, then it is always possible to find integers x and y not both zero such that

$$ax - by \equiv 0 \pmod{m} \tag{1}$$

and $|x| < m^{1/2}$, $|y| < m^{1/2}$.

In 1917 A. Thue proved [**10**]:

If a, b and m are relatively prime, then (1) *can be solved by integers x and y such that* $|x| \leqslant m^{1/2}$, $|y| \leqslant m^{1/2}$.

This is called, in general, the Theorem of Thue. See, for instance, the books of A. Scholz [**7**, p. 45], and O. Ore [**5**, p. 268]. If $(b, m) = 1$ and m is not a square, the results of Aubry and Thue are identical. If m is a square but $b/m^{1/2}$ is not an integer, then Aubry's result is better than Thue's. Since Aubry published the theorem first and Thue proved it independently a little later, it should be called the Theorem of Aubry-Thue. In addition to the books mentioned above, this theorem is also proved in the book of Uspensky and Heaslet [**11**, p. 234] without mentioning either Aubry or Thue.

Actually Thue had already proved in 1915 [**9**] a more general result under certain unimportant restrictions without formulating it as a theorem. If we omit these restrictions, Thue's result can be formulated as follows:

If $a_1, a_2, \ldots, a_n$ are relatively prime, then it is possible to find integers $x_1, x_2, \ldots, x_n$ not all zero such that

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n \equiv 0 \pmod{m} \tag{2}$$

and $0 \leqslant |x_\nu| \leqslant m^{1/n}$.

In 1926 J. M. Vinogradov [**12**] *generalized* the Theorem of Aubry-Thue in another direction:

Let p be a prime $(a, p) = 1$ and k any positive integer. Then there exist relatively prime integers x and y satisfying

$$ax \equiv y \pmod{p},\ 0 < x \leqslant k,\ 0 < |y| < p/k.$$

It is clear that the corresponding theorem holds for $ax \equiv by \pmod{p}$ where $(b, p) = 1$. Moreover it follows from the proof that the modulus need not be a prime number. In the book of Scholz [**7**] this generalization of Thue is also proved.

Received February 24, 1950.

Independently Thue's Theorem was proved by De Backer [3] and Vinogradov's generalization by Ballieu [2]. Moreover Ballieu considered the case where a and m are not relatively prime, but ordinarily it will be sufficient for the applications to consider the case where $(a, m) = 1$ since by must be divisible by the g. c. d. of a and m. In a second paper De Backer [4] stated without proof the following theorem which is unfortunately incorrect:

If $(a, m) = 1$ and if A is any integer, then

$$ax \equiv y + A \pmod{m}$$

always has a solution for which $|x| \leqslant m^{\frac{1}{2}}$, $|y| \leqslant m^{\frac{1}{2}}$.

For instance, $2x \equiv y + 23 \pmod{47}$ has no solution. De Backer used this result to prove the following theorem:

If p is a prime and a, b, c, d are integers, then the system

$$\begin{aligned} ax + by &\equiv z \pmod{p} \\ cx + dy &\equiv u \pmod{p} \end{aligned}$$

always has a solution x, y, z, u where each is less than $p^{\frac{1}{2}}$ in absolute value.

We wish to prove the latter theorem is correct by proving the following generalization of the theorem of Aubry-Thue which also contains (2) as a special case.

The system of r linear homogeneous congruences in s unknowns $(r < s)$

$$\sum_{\sigma=1}^{s} a_{\rho\sigma} x_\sigma \equiv 0 \pmod{m} \qquad (\rho = 1, 2, \ldots, r)$$

always has a non-trivial solution for which

$$|x_\sigma| \leqslant m^{r/s} \qquad (\sigma = 1, 2, \ldots, s).$$

This result will be obtained by proving the corresponding generalization of Vinogradov's theorem.

The theorem of Aubry-Thue is used in particular for the proof of the representation of primes of form $4n + 1$ as sum of two squares and that the least kth power non-residue (mod p) with $p \equiv 1 \pmod{k}$ is less than $p^{\frac{1}{2}}$. Correspondingly, we shall use our generalisation to simplify the proof that every integer can be represented as sum of four squares and we shall prove here that for odd k and $p \equiv 1 \pmod{k}$ each of the $k - 1$ classes of kth power non-residues contains at least one element less than $p^{(k-1)/k}$. For sufficiently large p and the special case $k = 3$ a sharper bound can be obtained from Vinogradov's results [13] but not for $k > 3$.

Porcelli and Pall have just announced that they can prove the following theorem with the help of Farey Series:

If p is an odd prime, D a quadratic residue (mod p), g and h positive integers such that $g \leqslant p$ and $h = [p/g]$, then at least one of the numbers $1^2, 2^2, \ldots, h^2$ is congruent to one of the numbers $D, 4D, 9D, \ldots, (g-1)^2D$.

We will show this theorem is an immediate consequence of Vinogradov's theorem and that it also holds for kth power residues with even k. Finally we will generalize the theorem of Aubry-Thue for congruences with regard to a double modulus and for congruences with respect to ideals in algebraic number fields.

2. Generalization of Vinogradov's Theorem.

THEOREM 1. *Let r and s be rational integers with $r < s$ and let f_σ be positive numbers less than m $(\sigma = 1, 2, \ldots, s)$ such that*

$$\prod_{\sigma=1}^{s} f_\sigma > m^r. \tag{3}$$

Then the system of r linear congruences

$$y_\rho = \sum_{\sigma=1}^{s} a_{\rho\sigma} x_\sigma \equiv 0 \pmod{m} \qquad (\rho = 1, 2, \ldots, r) \tag{4}$$

has a non-trivial solution in integers $x_1, x_2, \ldots, x_s$ such that

$$|x_\sigma| < f_\sigma \qquad (\sigma = 1, 2, \ldots, s). \tag{5}$$

Proof. Let f^*_σ be the greatest integer less than f_σ. For $\sigma = 1, 2, \ldots, s$ we choose

$$x_\sigma = 0, 1, 2, \ldots, f^*_\sigma \tag{6}$$

and obtain $\prod_{\sigma=1}^{s} (f^*_\sigma + 1)$ sets of r-tuples $(y_1, y_2, \ldots, y_r)$. By (3) we have

$$\prod_{\sigma=1}^{s} (f^*_\sigma + 1) \geqslant \prod_{\sigma=1}^{s} f_\sigma > m^r.$$

Thus it follows from Dirichlet's principle of the drawers that at least two of the r-tuples, say $(y'_1, y'_2, \ldots, y'_r)$ and $(y''_1, y''_2, \ldots, y''_r)$ will satisfy the congruences

$$y'_\rho \equiv y''_\rho \pmod{m} \qquad (\rho = 1, 2, \ldots, r). \tag{7}$$

If we denote the corresponding values of x_σ by x'_σ and x''_σ respectively, we have

$$\begin{aligned} y'_\rho &= a_{\rho 1}x'_1 + a_{\rho 2}x'_2 + \ldots + a_{\rho s}x'_s \\ y''_\rho &= a_{\rho 1}x''_1 + a_{\rho 2}x''_2 + \ldots + a_{\rho s}x''_s \end{aligned} \qquad (\rho = 1, \ldots, r).$$

Hence by (7) for $\rho = 1, 2, \ldots, r$,

$$a_{\rho 1}(x'_1 - x''_1) + a_{\rho 2}(x'_2 - x''_2) + \ldots + a_{\rho s}(x'_s - x''_s) \equiv 0 \pmod{m}.$$

If we denote $x'_\sigma - x''_\sigma$ by X_σ, then $X_1, X_2, \ldots, X_s$ are a non-trivial solution of the congruences (4) which by (6) satisfy the conditions (5).

COROLLARY 1. *Let $f(x)$ be any irreducible monic polynomial of degree n and p any prime. Let $f_1, f_2, \ldots, f_{2n}$ be positive numbers less than p such that*

$$\prod_{\nu=1}^{2n} f_\nu > p^n.$$

If $g(x)$ and $h(x)$ are given polynomials with integral rational coefficients, then we can find polynomials with integral rational coefficients:

$$\phi(x) = u_1x^{n-1} + u_2x^{n-2} + \ldots + u_n,$$
$$\psi(x) = v_1x^{n-1} + v_2x^{n-2} + \ldots + v_n,$$

not both zero such that,

$$g(x)\phi(x) + h(x)\psi(x) \equiv 0 \pmod{f(x), p}, \tag{8}$$

where

$$|u_\nu| < f_\nu, \qquad |v_\nu| < f_{\nu+n} \qquad (\nu = 1, 2, \ldots, n).$$

Proof. The coefficients of $g(x)\phi(x)$ are linear forms in $u_1, u_2, \ldots, u_n$. If we divide $g(x)\phi(x)$ by the monic polynomial $f(x)$, then the coefficients of the remainder are also linear combinations of $u_1, u_2, \ldots, u_n$ with given integral rational coefficients. Similarly the remainder of $h(x)\psi(x)$ after dividing by $f(x)$ will have coefficients which are linear forms in $v_1, v_2, \ldots, v_n$ with given integral rational coefficients. In order that (8) may hold, at most n linear congruences in the $2n$ variables u_ν and v_ν must be satisfied. Hence the corollary follows at once from Theorem 1.

If $A = (a_{\rho\sigma})$ and $B = (b_{\rho\sigma})$ are matrices with integral elements, we write $A \equiv B \pmod m$ if $a_{\rho\sigma} \equiv b_{\rho\sigma} \pmod m$ for every ρ and σ.

COROLLARY 2. *Let $f_{\sigma\tau}$ and $f'_{\sigma\tau}$ be positive numbers less than m ($\sigma = 1, 2, \ldots, s$; $\tau = 1, 2, \ldots, t$) such that*

$$\prod f_{\sigma\tau} f'_{\sigma\tau} > m^{rt}.$$

Let $A = (a_{\rho\sigma})$ and $B = (b_{\rho\sigma})$ be two $r \times s$ matrices with integral rational elements and $r < 2s$. Then for every given integer t we can find integral $s \times t$ matrices $U = (u_{\sigma\tau})$ and $V = (v_{\sigma\tau})$ such that

$$AU \equiv BV \pmod m, \tag{9}$$

where

$$|u_{\sigma\tau}| < f_{\sigma\tau}, |v_{\sigma\tau}| < f'_{\sigma\tau}, \quad (\sigma=1,2,\ldots,s; \tau=1,2,\ldots,t).$$

Proof. The rt elements of AU are linear combinations of the elements of U. Hence (9) requires that rt linear congruences for the $2st$ unknown elements of U and V be satisfied.

A similar result holds for left-hand multiplication of A and B.

3. The Four Square Theorem. It is well known that it is sufficient to prove this theorem only for prime numbers p. The simplest proofs use the fact that we can find integers a and b such that

$$a^2 + b^2 + 1 \equiv 0 \pmod p \tag{10}$$

and the method of descent [11, pp. 383-6]. We wish to prove that the theorem follows easily from (10) and Theorem 1.

Let a and b satisfy (10), then the congruences

$$\begin{aligned} x &\equiv az + bt \pmod p \\ y &\equiv bz - at \pmod p \end{aligned} \tag{11}$$

have a non-trivial solution with

$$\max(|x|, |y|, |z|, |t|) < p^{\frac{1}{2}}. \tag{12}$$

It follows from (11) and (10) that

$$x^2 + y^2 \equiv (a^2 + b^2)z^2 + (a^2 + b^2)t^2 \equiv -z^2 - t^2 \pmod p.$$

Hence

$$x^2 + y^2 + z^2 + t^2 = Ap. \tag{13}$$

By (12), A must be equal to 1, 2, or 3. If $A = 1$, the theorem is proved. If $A = 2$, then x must be congruent (mod 2) to at least one of y, z, t say $x \equiv y$ (mod 2) and then also $z \equiv t$ (mod 2). We obtain from (13) for p the following representation as sum of four squares:

$$p = \left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 + \left(\frac{z+t}{2}\right)^2 + \left(\frac{z-t}{2}\right)^2.$$

If $A = 3$, we use a method of Sylvester [8]. It follows from (13) that one of x, y, z, t, say x, must be divisible by 3 and by proper choice of signs for y, z, and t we may assume that

$$y \equiv z \equiv t \pmod 3.$$

Hence from (13)

$$p = \left(\frac{y+z+t}{3}\right)^2 + \left(\frac{x+z-t}{3}\right)^2 + \left(\frac{x-y+t}{3}\right)^2 + \left(\frac{x+y-z}{3}\right)^2.$$

This gives our representation since the parentheses are integers and hence proves our theorem.

4. The least kth power non-residues.

THEOREM 2. *If k is odd and p a prime where $p \equiv 1 (\mathrm{mod}\ k)$, then each of the $k - 1$ classes of kth power non-residues contains at least one element which is less than $p^{(k-1)/k}$.*

Proof. Let $n_1, n_2, \ldots, n_{k-1}$ be representatives of the $k-1$ classes K_1, $K_2, \ldots, K_{k-1}$ of non-residues. We consider the system of $k-1$ congruences in k unknowns:

$$
\begin{aligned}
x &\equiv n_1 y_1 \pmod p \\
x &\equiv n_2 y_2 \pmod p \qquad (14)\\
&\cdots \\
x &\equiv n_{k-1} y_{k-1} \pmod p.
\end{aligned}
$$

This system has a non-trivial solution $x, y_1, y_2, \ldots, y_{k-1}$ where each unknown is less than $p^{(k-1)/k}$ in absolute value. Since -1 is a kth power residue for odd k, then x and $-x$ belong to the same class. Hence we only have to show that $x, y_1, y_2, \ldots, y_{k-1}$ are representatives of the k classes of residues and non-residues. If x belongs to the class K of residues or non-residues, then y_i belongs to the class $K K_i^{-1}$ $(i = 1, 2, \ldots, k-1)$. It is obvious that these classes are different from each other and different from K.

If we consider instead of the $k-1$ congruences (14) only l of them, then it follows in the same way from Theorem 1 that l of these classes of kth power non-residues contain elements which are less than $p^{l/l+1}$. Applying this successively for $l = 1, 2, \ldots, k-1$, we obtain

THEOREM 3. *If k is odd and p a prime with $p \equiv 1$ (mod k), then it is possible to find $k-1$ non-residues $d_1, d_2, \ldots, d_{k-1}$ belonging to different classes such that*

$$0 < d_\lambda < p^{\lambda/(\lambda+1)}, \qquad (\lambda = 1, 2, \ldots, k-1).$$

This gives for d_1 the well known bound for the least kth power non-residue.

5. Generalization of a Theorem of Porcelli and Pall.

THEOREM 4. *Let g and k be positive integers where k is even, p an odd prime with $p \equiv 1$ (mod k) such that $g \leqslant p$. We set $h = [p/g]$. If D is a kth power residue, then at least one of the numbers $1^k, 2^k, \ldots, h^k$ is congruent to one of the numbers $D, 2^k D, \ldots, (g-1)^k D$.*

Proof. Since D is a kth power residue, there exists an integer a such that $a^k \equiv D \pmod p$. By Theorem 1, the congruence

$$ax \equiv y \pmod p$$

has a solution for which $|x| < g$ and $|y| < h + 1$ since $g(h+1) > p$. Thus

$$a^k x^k \equiv y^k \pmod p$$

and since k is even,

$$D\,|x|^k \equiv |y|^k \pmod p.$$

Since $|x|$ is one of the numbers $1, 2, \ldots, g-1$ and $|y|$ one of the numbers $1, 2, \ldots, h$, the theorem is proved.

6. A generalization for algebraic numbers.

THEOREM 5. *Let* $\mathfrak{m}$ *be an ideal of an algebraic number field and* t *the norm of* $\mathfrak{m}$. *Assume that* t *is less than the square of the smallest rational integer* g *of* $\mathfrak{m}$. *If* α *and* β *are two integers of the field, then the congruence*

$$\alpha x - \beta y \equiv 0 \pmod{\mathfrak{m}} \tag{15}$$

has a solution in rational integers x *and* y *not both belonging to* $\mathfrak{m}$ *such that*

$$|x| < t^{1/2}, \quad |y| < t^{1/2}. \tag{16}$$

Proof. The numbers $0, 1, 2, \ldots, [t^{1/2}]$ are incongruent (mod $\mathfrak{m}$) since their difference is less than g. If we choose for x and y the numbers $0, 1, 2, \ldots, [t^{1/2}]$, then we obtain $\{[t^{1/2}] + 1\}^2$ numbers $\alpha x - \beta y$, hence more than t integers of the field. At least two of them, say $\alpha x' - \beta y'$ and $\alpha x'' - \beta y''$ must be congruent (mod $\mathfrak{m}$). Hence

$$\alpha(x' - x'') - \beta(y' - y'') \equiv 0 \pmod{\mathfrak{m}}$$

and $X = x' - x''$, $Y = y' - y''$ are a solution of (15) satisfying (16), $x' - x'' \equiv 0 \pmod{\mathfrak{m}}$ implies $x' = x''$ and $y' - y'' \equiv 0 \pmod{\mathfrak{m}}$ implies $y' = y''$.

The assumptions of Theorem 5 are satisfied, for instance, if $\mathfrak{m}$ is the product of different prime ideals of degree 1 of which no two are conjugates. If, namely,

$$\mathfrak{m} = \mathfrak{p}_1 \mathfrak{p}_2 \ldots \mathfrak{p}_t$$

and $p_1, p_2, \ldots, p_t$ the prime numbers contained in these ideals, then $p_1, p_2, \ldots, p_t$ are different and

$$t = p_1 . p_2 \ldots p_t.$$

On the other hand, $p_1 p_2 \ldots p_t$ is the smallest positive integer contained in $\mathfrak{m}$. The theorem holds also if some of these prime ideals but not all are of degree 2.

Note (May 4, 1951). In the meantime the paper of Porcelli and Pall has been published [6]. While in their abstract only the case $k = 2$ is mentioned. actually Theorem 4 is proved in the paper. Our proof is completely different from the proof of Porcelli and Pall.

REFERENCES

1. L. Aubry, *Un théorème d'arithmétique*, Mathesis (4), vol. 3 (1913).

2. R. Ballieu, *Sur des congruences arithmétiques*, Bulletin de la Classe des Sciences de l'Académie Royale de Belgique (5), vol. 34 (1948), 39-45.

3. S. M. De Backer, *Un théorème fundamental*, Bulletin de la Classe des Sciences de l'Académie Royale de Belgique (5), vol. 33 (1947) 632-634.

4. ——— *Solutions modérées d'un système de congruences du premier degré pour un module premier p*. Bulletin de la Classe des Sciences de l'Academie Royale de Belgique, (5) vol. 34 (1948), 46-51.

5. O. Ore, *Number theory and its history* (New York, 1948), 268.

6. P. Porcelli and G. Pall, *A property of Farey sequences*, Can. J. Math., vol. 3 (1951) 52-53.

7. A. Scholz, *Einführung in die Zahlentheorie* (Berlin, 1939).

8. J. J. Sylvester, *Note on a principle in the theory of numbers and the resolubility of any number into the sum of four squares*, Quar. J. of Math., vol. 1 (1857), 196-7; or Collected Math. Papers, vol. 2 (1908), 101-102.

9. A. Thue, *Über die ganzzahlige Gleichung* $C^n = a^m + a^{m-1}b + \ldots + ab^{m-1} + b^m$, Norske videnskaps-akademi, Oslo, Matematisk-naturvidenskapelig klasse Skrifter, No. 3 (1915).

10. ——— *Et bevis for at lignigen* $A^3 + B^3 = C^3$ *er remulig i hele fra nul forsk jellige tal* A, B, *og* B.Archiv. for Math. og Naturvid, vol. 34, No. 15 (1917).

11. J. V. Uspensky and M. A. Heaslit, *Elementary number theory* (New York, 1939).

12. J. M. Vinogradov, *On a general theorem concerning the distribution of the residues and non-residues of powers*, Trans. Amer. Math. Soc., vol. 29 (1927), 209-17.

13. ——— *On the bound of the least non-residues of nth powers*, Trans. Amer. Math. Soc. vol. 29 (1927), 218-226.

University of North Carolina

ON THE CHANGES OF SIGN OF A CERTAIN ERROR FUNCTION

PAUL ERDÖS AND HAROLD N. SHAPIRO

1. Introduction. Though much effort has been expended in studying the mean values of arithmetic functions there is one case which has not yielded a great deal either to elementary or analytic methods. The case to which we refer is that of estimating

$$\Phi(x) = \sum_{n \leq x} \phi(n), \tag{1.1}$$

where $\phi(n)$ is the Euler function (i.e. $\phi(n)$ = the number of integers less than n which are relatively prime to n). If we define the error function $R(x)$ via

$$R(x) = \Phi(x) - \frac{3}{\pi^2} x^2, \tag{1.2}$$

the question reduces to studying the behaviour of $R(x)$. The first result is due to Dirichlet [1], who proved that

$$R(x) = O(x^\delta) \tag{1.3}$$

for some δ, $1 < \delta < 2$. This was improved by Mertens [2] to

$$R(x) = O(x \log x). \tag{1.4}$$

The proofs in both cases are very short and simple and may be found in various textbooks [1], [3]. It is therefore of particular interest that to date there has been no improvement in the estimate for $R(x)$ beyond (1.4).

In a different direction Pillai and Chowla [4] have proved that

$$R(x) \neq o(x \log \log \log x), \tag{1.5}$$

and

$$\sum_{n \leq x} R(n) \sim \frac{3}{2\pi^2} x^2. \tag{1.6}$$

Sylvester, [5], [6], conjectured among other things that for all integers $x > 0$, $R(x) > 0$. This was disproved by M. L. N. Sarma [7], by the simple expedient of showing that $R(820) < 0$.

In this paper we propose to prove that $R(x)$ changes sign for infinitely many integers x. More precisely, there exists a positive constant c and infinitely many integers x such that

$$R(x) > c\, x \log \log \log \log x, \tag{1.7}$$

and infinitely many integers x such that

Received July 8, 1950.

$$R(x) < -c\,x \log\log\log\log x. \tag{1.8}$$

2. The evaluation of certain sums. The proofs of the results mentioned in the introduction are obtained by first treating the error function

$$H(x) = \sum_{n\leqslant x} \frac{\phi(n)}{n} - \frac{6}{\pi^2}x.$$

The relationship between $H(x)$ and $R(x)$ is given by

LEMMA 2.1. *For integral* x,

$$\sum_{n\leqslant x} H(n) = \frac{3}{\pi^2}x + (x+1)H(x) - R(x). \tag{2.1}$$

Proof.

$$\begin{aligned}\sum_{n\leqslant x} H(n) &= \sum_{n\leqslant x}\left\{\sum_{m\leqslant n}\frac{\phi(m)}{m} - \frac{6}{\pi^2}n\right\}\\ &= \sum_{n\leqslant x}(x-n+1)\frac{\phi(n)}{n} - \frac{3}{\pi^2}x(x+1)\\ &= (x+1)\left\{\frac{6}{\pi^2}x + H(x)\right\} - \sum_{n\leqslant x}\phi(n) - \frac{3}{\pi^2}x(x+1)\\ &= \frac{3}{\pi^2}x + (x+1)H(x) - R(x).\end{aligned}$$

We will need estimates for certain sums which we now provide.

LEMMA 2.2.

$$\sum_{d\leqslant x}\frac{1}{d}H\left(\frac{x}{d}\right) = O(1), \tag{2.2}$$

$$\sum_{d\leqslant x} H\left(\frac{x}{d}\right) = O(x), \tag{2.3}$$

$$\sum_{d\leqslant x} R\left(\frac{x}{d}\right) = O(x). \tag{2.4}$$

Proof. (2.3) follows immediately from the fact that $H(x) = O(\log x)$. Next we consider (2.2):

$$\begin{aligned}x + O(1) = \sum_{n\leqslant x} 1 &= \sum_{n\leqslant x}\frac{1}{n}\sum_{d\mid n}\phi(d)\\ &= \sum_{dd'\leqslant x}\frac{\phi(d)}{dd'} = \sum_{d\leqslant x}\frac{1}{d}\left\{\frac{6}{\pi^2}\frac{x}{d} + H\left(\frac{x}{d}\right)\right\}\\ &= x + O(1) + \sum_{d\leqslant x}\frac{1}{d}H\left(\frac{x}{d}\right).\end{aligned}$$

which yields (2.2). Similarly,

$$\frac{x^2}{2} + O(x) = \sum_{n \leqslant x} n = \sum_{n \leqslant x} \sum_{d|n} \phi(d)$$

$$= \sum_{d \leqslant x} \sum_{d' \leqslant x|d} \phi(d)' = \sum_{d \leqslant x} \left\{ \frac{3}{\pi^2} \frac{x^2}{d^2} + R\left(\frac{x}{d}\right) \right\};$$

whence (2.4) follows.

THEOREM 2.1.

$$\sum_{mn \leqslant x} H(n) = \frac{3}{\pi^2} x \log x + O(x). \tag{2.5}$$

Proof. From Lemma 2.1 we obtain for all $x > 0$ that

$$\sum_{n \leqslant x} H(n) = \frac{3}{\pi^2} x + x\, H(x) - R(x) + O(\log x). \tag{2.6}$$

Replacing x by x/m in (2.5) and summing over all integral $m \leqslant x$ we have

$$\sum_{m \leqslant x} \sum_{n \leqslant x|m} H(n) = \frac{3}{\pi^2} \sum_{m \leqslant x} \frac{x}{m} + x \sum_{m \leqslant x} \frac{1}{m} H\left(\frac{x}{m}\right) - \sum_{m \leqslant x} R\left(\frac{x}{m}\right) + O\ (\log x).$$

Then, taking into account the estimates of Lemma 2.2 we obtain (2.5).

Actually, Pillai and Chowla [4] have proved that

$$\sum_{n \leqslant x} H(n) \sim \frac{3}{\pi^2} x, \tag{2.7}$$

and we could use (2.7) instead of (2.5) in our development. However, the proof of (2.7) requires the prime number theorem, and we therefore introduce (2.5) for the sake of simplicity.

3. The average of $H(n)$ over arithmetic progressions. The main part of our proof consists of evaluating certain averages of $H(n)$ over arithmetic progressions. We begin with

LEMMA 3.1.

$$\sum_{\substack{m \leqslant z \\ m \equiv \beta(A)}} \frac{\phi(m)}{m} = \frac{C}{A} \sum_{d|(A,\beta)} \frac{\mu(d)}{d} z + O(\log z), \tag{3.1}$$

where

$$C = C(A) = \prod_{p \nmid A} \left(1 - \frac{1}{p^2}\right).$$

Proof.

$$\sum_{\substack{m \leqslant z \\ m \equiv \beta(A)}} \frac{\phi(m)}{m} = \sum_{\substack{dd' \equiv \beta(A) \\ dd' \leqslant z}} \frac{\mu(d)}{d}$$

$$= \sum_{\substack{d \leqslant z \\ (d, A)|\beta}} \frac{\mu(d)}{d} \left\{ \frac{(d,A)}{A} \frac{z}{d} + O(1) \right\}$$

$$= \frac{z}{A} \sum_{\tau | (A,B)} \tau \sum_{(dA)=\tau} \frac{\mu(d)}{d^2} + O(\log z)$$

$$= \frac{z}{A} \sum_{\tau | (A,B)} \tau \frac{\mu(\tau)}{\tau} \sum_{(t,A)=1} \frac{\mu(t)}{t^2} + O(\log z)$$

$$= \frac{Cz}{A} \sum_{\tau | (A,B)} \frac{\mu(\tau)}{\tau} + O(\log z).$$

THEOREM 3.1. *For A, B any integers such that $A > B \geqslant 0$*

$$\sum_{n \leqslant x} H(An - B) = \frac{1}{A} \sum_{n \leqslant Ax} H(n) + \Lambda x + O(\log x) \tag{3.2}$$

where

$$\Lambda = \Lambda(A, B) = M(A, B) - 3/\pi^2,$$

$$M(A,B) = \begin{cases} \dfrac{6}{\pi^2} B - \dfrac{1}{2} \dfrac{\phi(A)C(A)}{A} - C(A) \displaystyle\sum_{c=1}^{B-1} \frac{\phi(A,c)}{(A,c)} & \text{for } B \neq 0; \\ \dfrac{1}{2} \dfrac{\phi(A)C(A)}{A} & \text{for } B = 0. \end{cases} \tag{3.3}$$

Proof. It clearly suffices to prove (3.2) for x integral, and so we assume x an integer. We have

$$\begin{aligned} \sum_{n \leqslant x} H(An - B) &= \sum_{n \leqslant x} \sum_{m \leqslant An - B} \frac{\phi(m)}{m} - \frac{6}{\pi^2} \sum_{n \leqslant x} (An - B) \\ &= \sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} \sum_{\frac{m+B}{A} \leqslant n \leqslant x} 1 - \frac{3}{\pi^2} (Ax^2 + Ax - 2Bx) \\ &= \sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} \left\{ x - \left[\frac{m+B}{A} \right] \right\} + \sum_{\substack{m \leqslant Ax - B \\ m \equiv -B(A)}} \frac{\phi(m)}{m} \\ &\qquad - \frac{3}{\pi^2} (Ax^2 + Ax - 2Bx). \end{aligned} \tag{3.4}$$

Considering the first sum of (3.4) we have

$$\sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} \left\{ x - \left[\frac{m+B}{A} \right] \right\} = x \sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} - \sum_{a=0}^{A-1} \sum_{\substack{m \leqslant Ax - B \\ m+B \equiv a(A)}} \frac{\phi(m)}{m} \frac{m+B-a}{A}$$

$$= \left\{ x \sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} - \frac{1}{A} \sum_{m \leqslant Ax - B} \phi(m) \right\} - \sum_{a=0}^{A-1} \frac{(B-a)}{A} \sum_{\substack{m \leqslant Ax - B \\ m+B \equiv a(A)}} \frac{\phi(m)}{m}$$

$$= \frac{1}{A} \left\{ (Ax - B + 1) \sum_{m \leqslant Ax - B} \frac{\phi(m)}{m} - \sum_{m \leqslant Ax - B} \phi(m) - \frac{3}{\pi^2} [(Ax - B)^2 + (Ax - B)] \right\}$$

$$(3.5)\quad \begin{aligned}&+\frac{B-1}{A}\sum_{m\leq Ax-B}\frac{\phi(m)}{m}+\sum_{a=0}^{A-1}\frac{(a-B)}{A}\sum_{\substack{m\leq Ax-B\\ m+B\equiv a(A)}}\frac{\phi(m)}{m}\\ &+\frac{3}{\pi^2}\frac{[(Ax-B)^2+(Ax-B)]}{A}\end{aligned}$$

$$=\frac{1}{A}\sum_{n\leq Ax-B}H(n)+\frac{3}{\pi^2}Ax^2-\frac{3}{\pi^2}x+\sum_{a=0}^{A-1}\frac{(a-B)}{A}\sum_{\substack{m\leq Ax-B\\ m+B\equiv a(A)}}\frac{\phi(m)}{m}+O(\log x).$$

Next, using Lemma 3.1, we note that

$$(3.6)\quad \sum_{a=0}^{A-1}\frac{(a-B)}{A}\sum_{\substack{m\leq Ax-B\\ m+B\equiv a(A)}}\frac{\phi(m)}{m}=\frac{x}{A}\sum_{a=0}^{A-1}(a-B)\,C(A)\sum_{d|(A,\,a-B)}\frac{\mu(d)}{d}+O(\log x).$$

On the other hand,

$$(3.7)\quad \begin{aligned}\sum_{a=0}^{A-1}(a-B)\sum_{d|(A,\,a-B)}\frac{\mu(d)}{d}&=\sum_{c=-B}^{A-B-1}c\sum_{d|(A,\,c)}\frac{\mu(d)}{d}\\ &=\sum_{c=0}^{A-B-1}c\sum_{d|(A,\,c)}\frac{\mu(d)}{d}+\sum_{c=A-B}^{A-1}(c-A)\sum_{d|(A,\,c)}\frac{\mu(d)}{d}\\ &=\sum_{c=0}^{A-1}c\sum_{d|(A,\,c)}\frac{\mu(d)}{d}-A\sum_{c=A-B}^{A-1}\sum_{d|(A,\,c)}\frac{\mu(d)}{d}\\ &=\sum_{c=0}^{A-1}c\sum_{d|(A,\,c)}\frac{\mu(d)}{d}-A\sum_{c=1}^{B}\sum_{d|(A,\,c)}\frac{\mu(d)}{d}.\end{aligned}$$

For each term of (3.7) we have in turn

$$(3.8)\quad \begin{aligned}\sum_{c=0}^{A-1}c\sum_{d|(A,\,c)}\frac{\mu(d)}{d}&=\sum_{d|A}\frac{\mu(d)}{d}\sum_{\substack{1\leq c\leq A-1\\ c\equiv 0(d)}}c\\ &=\frac{1}{2}\sum_{d|A}\mu(d)\left\{\left(\frac{A}{d}\right)^2-\left(\frac{A}{d}\right)\right\}\\ &=\frac{1}{2}A^2\sum_{d|A}\frac{\mu(d)}{d^2}-\frac{1}{2}\phi(A);\end{aligned}$$

and

$$(3.9)\quad \sum_{c=1}^{B}\sum_{d|(A,\,c)}\frac{\mu(d)}{d}=\sum_{c=1}^{B}\frac{\phi(A,c)}{(A,c)},$$

where this last sum is 0 if $B=0$.

Combining (3.6), (3.7), (3.8), and (3.9) we get

$$(3.10)\quad \begin{aligned}\sum_{a=0}^{A-1}\left(\frac{a-B}{A}\right)\sum_{\substack{m\leq Ax-B\\ m+B\equiv a(A)}}\frac{\phi(m)}{m}=x\Bigg\{&\frac{C(A)A}{2}\sum_{d|A}\frac{\mu(d)}{d^2}-\frac{1}{2}\frac{\phi(A)C(A)}{A}\\ &-C(A)\sum_{c=1}^{B}\frac{\phi(A,c)}{(A,c)}\Bigg\}+O(\log x).\end{aligned}$$

Finally, inserting this in (3.5), noting that $C(A) \sum_{d|A} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2}$, and combining with (3.4) and Lemma 3.1 we obtain

$$\begin{aligned} \sum_{m\leqslant Ax} H(An-B) &= \frac{1}{A} \sum_{n\leqslant Ax} H(n) + \frac{3}{\pi^2} Ax^2 - \frac{3}{\pi^2} x \\ &\quad + \frac{3}{\pi^2} Ax - \frac{x}{2} \frac{\phi(A)C(A)}{A} - C(A)x \sum_{c=1}^{B} \frac{\phi(A,c)}{(A,c)} \\ &\quad + \frac{C(A)x\phi(A,B)}{(A,B)} - \frac{3}{\pi^2}[Ax^2+Ax-2Bx]+O(\log x) \\ &= \frac{1}{A} \sum_{n\leqslant Ax} H(n) + \Lambda x + O(\log x). \end{aligned} \tag{3.11}$$

THEOREM 3.2. *For A, B any integers, $A > B \geqslant 0$,*

$$\sum_{mn\leqslant x} H(An - B) = M(A,B)\, x \log x + O(x). \tag{3.12}$$

Proof. Replacing x by x/m in (3.2) and summing over all integers $m \leqslant x$, we have

$$\sum_{mn\leqslant x} H(An-B) = \frac{1}{A} \sum_{m\leqslant x} \sum_{n\leqslant Ax|m} H(n) + \Lambda x \log x + O(x).$$

Since

$$\sum_{x<m\leqslant Ax} \sum_{n\leqslant Ax|m} H(n) = O\left(\sum_{m\leqslant Ax} 1 \right) = O(x),$$

we get

$$\sum_{mn\leqslant x} H(An-B) = \frac{1}{A} \sum_{mn\leqslant Ax} H(n) + \Lambda\, x \log x + O(x), \tag{3.13}$$

so that via (2.5) this reduces to (3.12).

We note in passing that if we combine (3.2) with the deeper result (2.7) we have

THEOREM 3.3. *For A, B any integers, $A > B \geqslant 0$,*

$$\sum_{n\leqslant x} H(An - B) \backsim M(A, B)x. \tag{3.14}$$

4. On the changes of sign of $H(x)$. Merely to show that $H(x)$ changes sign infinitely often is easily deduced from (3.12). We note first that if $A = A_\kappa = \prod_{i=1}^{\kappa} p_i$, and κ is sufficiently large

$$\sum_{c=1}^{B-1} \frac{\phi(A,c)}{(A,c)} = \sum_{c=1}^{B-1} \frac{\phi(c)}{c} = \frac{6}{\pi^2}(B-1) + H(B-1).$$

Thus we obtain easily for $B \neq 0$, and fixed, that

$$\lim_{\kappa\to\infty} \lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_\kappa n - B) = \frac{6}{\pi^2} - H(B-1).$$

Since

$$\frac{6}{\pi^2} - H(B-1) = \frac{\phi(B)}{B} - H(B),$$

this may be written as

$$\lim_{\kappa\to\infty} \lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_\kappa n - B) = \frac{\phi(B)}{B} - H(B). \tag{4.1}$$

From (2.5) it follows that $H(n)$ is positive for infinitely many n, and we need only show that we cannot have $H(n) \geqslant 0$ for all sufficiently large n. For if this were so, for all sufficiently large B

$$\lim_{\kappa\to\infty} \lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_\kappa n - B) \geqslant 0,$$

so that we would have

$$\frac{\phi(B)}{B} \geqslant H(B) \geqslant 0.$$

For $\epsilon > 0$, small, choosing a large odd number B such that $\dfrac{\phi(B)}{B} < \epsilon$, we see that

$$H(B+1) = H(B) - \frac{6}{\pi^2} + \frac{\phi(B+1)}{B+1} \leqslant \epsilon - \frac{6}{\pi^2} + \frac{1}{2} < 0,$$

which would provide a contradiction.

The above argument can be improved upon if we use the analogue of (1.5) for $H(x)$ in conjunction with (4.1). This analogue, also proved by Pillai and Chowla, asserts that

$$H(x) \neq o(\log\log\log x). \tag{4.2}$$

Thus their exist infinitely many integral x such that

$$|H(x)| > c \log\log\log x, \tag{4.3}$$

where k is some positive constant. From (4.3) we note that given any large number $N \geqslant 6$ we can find an integer B such that $|H(B)| > N$. We then examine two cases:

Case 1. $H(B) > N$.

In this case we obtain from (4.1) that

$$\lim_{k\to\infty} \lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_k n - B) < -N + 1;$$

and for all sufficiently large k, say $k \geqslant k_0$, we have

$$\lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_k n - B) < -N + 2.$$

Then for each such k there exists an $x_0 = x_0(k)$ such that, for all $x \geqslant x_0$,

$$\sum_{mn \leqslant x} H(A_k n - B) < (-N + 3)\, x \log x; \tag{4.4}$$

from (4.4) we see that for each $k \geqslant k_0$ we obtain an $n^* = n^*(k)$ such that

$$H(A_k n^* - B) < -N + 3 \leqslant -\tfrac{1}{2}N.$$

Case 2. $H(B) < -N$.

In this case we proceed exactly as in Case (1), obtaining from (4.1) that

$$\lim_{k\to\infty} \lim_{x\to\infty} \frac{1}{x \log x} \sum_{mn \leqslant x} H(A_k n - B) > N.$$

This in turn yields a k_0 such that for each $k \geqslant k_0$ there is an $n^* = n^*(k)$ such that

$$H(A_k n^* - B) \geqslant \tfrac{1}{2}N.$$

From the above we see that $H(x)$ assumes arbitrarily large positive and negative values. We may restate this and its implication for $R(x)$ as follows.

THEOREM 4.1. *For integral x, we have*

$$\overline{\lim}\, H(x) = \infty \quad \text{and} \quad \underline{\lim}\, H(x) = -\infty, \tag{4.5}$$

$$\overline{\lim}\, \frac{R(x)}{x} = \infty \quad \text{and} \quad \underline{\lim}\, \frac{R(x)}{x} = -\infty. \tag{4.6}$$

Proof. (4.5) is clear from the above remarks. From (2.1) and (2.7) (or the weaker estimate $\sum_{n \leqslant x} H(n) = O(x)$), we obtain

$$R(x) = x\, H(x) + O(x), \tag{4.7}$$

and (4.6) then follows from (4.5).

5. More precise results. By refining some of our estimates the arguments used above may be made to yield the still more precise result that for some $c > 0$, there exist infinitely many integers x such that

$$H(x) > c \log \log \log \log x, \tag{5.1}$$

and infinitely many such that

$$H(x) < -c \log \log \log \log x. \tag{5.2}$$

We shall now give a sketch of the proof of this.

We need to obtain the dependence of many of the estimates obtained above on the modulus A. To begin with, a glance at the proof of Lemma 3.1 yields

$$\sum_{\substack{m \leqslant z \\ m \equiv B(A)}} \frac{\phi(m)}{m} = \frac{C}{A} \sum_{d \mid (A,B)} \frac{\mu(d)}{d} z + O\left(\sum_{\tau \mid (A,B)} \mu^2(\tau) \log \frac{z}{\tau} \right). \tag{5.3}$$

Using (5.3) instead of (3.1) in the proof of Theorem 3.1 we obtain for integral x,

$$\sum_{n \leqslant x} H(An - B) = \frac{1}{A} \sum_{n \leqslant Ax} \Lambda(n) + \Lambda x + O(2^{\nu(A)} \log Ax), \tag{5.4}$$

where $\nu(A)$ = the number of distinct prime factors of A.

Combining (5.4) and (2.7) gives

$$\sum_{n \leqslant x} H(An - B) = M(A,B)x + O(2^{\nu(A)} \log Ax) + o(x), \tag{5.5}$$

where both the O and o are uniform in A. Then taking $x = A = \prod_{p \leqslant B} p$ and noting that then $1 - \frac{c_1}{B} < C(A) < 1 - \frac{c_2}{B}$ $(c_1 > 0,\ c_2 > 0)$, we obtain (for all sufficiently large B) that there is a constant l, independent of both A and B, such that

$$\left| \frac{1}{A} \sum_{n \leqslant A} H(An - B) + H(B) \right| \leqslant l. \tag{5.6}$$

The desired result now follows from (5.6). We know that for infinitely many B,

$$|H(B)| > c \log\log\log B.$$

There are then, two cases:

Case (a). $H(B) > k \log\log\log B$.

In this case (5.6) implies that there exists an $n^* \leqslant A$ such that

$$\begin{aligned} H(An^* - B) &\leqslant l - c \log\log\log B \\ &\leqslant -\tfrac{1}{2}c \log\log\log B \\ &\leqslant -c_1 \log\log\log\log (An^* - B), \end{aligned}$$

for large B, since for $A = \prod_{p \leqslant B} p$, $\log A \backsim B$.

Case (b). $H(B) < -c \log\log\log B$.

Then as in Case (a), (5.6) implies that there exists an $n^* \leqslant A$ such that

$$\begin{aligned} H(An^* - B) &\geqslant c \log\log\log B - l \\ &\geqslant \tfrac{1}{2}c \log\log\log B \\ &\geqslant c_1 \log\log\log\log (An^* - B). \end{aligned}$$

Thus we see that there exist infinitely many integers x such that each of the inequalities (5.1), (5.2) hold. Combining this information with (4.7) we obtain the analogous result for the inequalities (1.7) and (1.8).

University of Aberdeen
and
New York University

Editor's Note: References for this paper were not available at time of going to press. They will appear in the following number of the Journal.